



RESULT 2  
ID W83213 standard; Protein; 237 AA.  
AC W83213:  
DT 09-FEB-1999 (first entry)  
DE Prostate-specific antigen protein hK3.  
KW pphk2; prostate detection; hK2; hK1; hK3; pphk2; tissue kallikrein;  
KW pphk2; prostate-specific glandular kallikrein protein; PSA; human;  
OS Homo sapiens.  
PN W09845795-A1.  
PD 22-OCT-1998.  
PF 09-APR-1998; U07027.  
PR 11-APR-1997; US-843076.  
PA (BAYU ) BAYLOR COLLEGE MEDICINE.  
PI (MAYO-) MAYO FOUNDATION.  
PI Slawin KM, Tindall DJ, Young CYF;  
DR WPI: 98-594592/50.  
PT Detection of human kallikrein 2 RNA - by reverse transcription and  
PT amplification by PCR, for detecting, monitoring and staging of  
PT prostate cancer.  
PS Disclosure: Fig 1: 90pp: English.  
CC The present invention describes a diagnostic method for detecting human  
CC kallikrein 2 (hK2) DNA. The method comprises: (a) contacting DNA obtained  
CC by reverse transcription (RT) of RNA from a human physiological sample  
CC which comprises cells suspected of containing hK2 RNA with at least 2  
CC oligonucleotides to amplify the DNA by PCR to yield amplified hK2 DNA,  
CC where the conditions amplify the DNA obtained by RT of RNA from at least  
CC one cell containing hK2 in a sample which comprises at least 107 to 109  
CC cells; and (b) detecting the presence of the amplified hK2 DNA. The  
CC method can be used for detecting, monitoring the progression of and  
CC pathologically staging prostate cancer. The present sequence represents  
CC prostate-specific antigen protein hK3.  
SQ Sequence 237 AA:

Query Match 100.0%; Score 183; DB 1; Length 237;  
Best Local Similarity 100.0%; Pred. No. 3.20e-13;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 139 EEFLEPKKLCQVDLHVSNVCAOV 163  
QY 1 EEFLEPKKLCQVDLHVSNVCAOV 25

RESULT 3  
ID R77098 standard; Protein; 237 AA.  
AC R77098:  
DT 16-JAN-1996 (first entry)  
DE Prostate-specific antigen.  
KW Prostate-specific antigen; prostate cancer; diagnosis.  
OS Homo sapiens.  
PN W09528498-A1.  
PD 26-OCT-1995.  
PF 14-APR-1995; U04680.  
PR 15-APR-1994; US-229391.  
PA (UYCO ) UNIV COLUMBIA NEW YORK.  
PI Buttyan R, Katz AE, Olsson CA, Ratfo A;  
DR WPI: 95-373812/48.  
DR N-PSDB: T04864.  
PT Method for enhancing prostate-specific antigen detection - provides  
PT sensitive means to identify early stages of prostate cancer.  
PS Disclosure: Page 69-70; 94pp: English.  
CC A prostate-specific antigen (PSA) is detected by isolating mRNA from  
CC a sample, producing cDNA, and subjecting the cDNA to RT-PCR using  
CC primers specific for the human PSA coding region (given in T04864).  
SQ Sequence 237 AA:

Query Match 100.0%; Score 183; DB 1; Length 237;  
Best Local Similarity 100.0%; Pred. No. 3.20e-13;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 139 EEFLEPKKLCQVDLHVSNVCAOV 163  
QY 1 EEFLEPKKLCQVDLHVSNVCAOV 25

RESULT 4  
ID W56086 standard; Protein; 237 AA.  
AC W56086:  
DT 11-AUG-1998 (first entry)  
DE Human prostate specific antigen protein.  
KW Human; prostate specific antigen; PSA; epitope; monoclonal antibody;  
KW detection; cancer; serine protease.  
OS Homo sapiens.  
PN W09810292-A1.  
PD 12-MAR-1998.  
PF 25-AUG-1997; U14909.  
PR 06-SEP-1996; US-025404.  
PA (GENZ ) CENTOCOR INC.  
PI Heavner GA;  
DR WPI: 98-193789/17.  
PT Monoclonal antibodies specific for prostate specific antigen -  
PT useful, e.g. in screening for prostate or breast cancer and  
PT especially to distinguish between benign prostatic hyperplasia and  
PT prostate cancer.  
PS Disclosure: Page 32-33; 84pp: English.  
CC The present sequence represents prostate specific antigen (PSA).  
CC Monoclonal antibodies specific for PSA and hydrolases producing them  
CC have been developed. The antibodies: (a) bind to free PSA; (b) are  
CC monoclonal antibody 365 binding to amino acids 82-87 of free PSA (LKNRFL)  
CC or fragment, and (c) are monoclonal antibodies 10, 11, 16 or 22.2 and  
CC 15.2, 156 or 225 binding to amino acids 139-144 (EEFLFIP) and 55-60  
CC (SLRHPF) respectively of free and bound PSA, or fragments. The antibodies  
CC are useful to detect PSA. For detecting free PSA only, an immunoassay  
CC comprising a solid support with attached labelled monoclonal antibody  
CC specific for free PSA (especially (b)) and a PSA standard can be used.  
CC To detect both free and bound PSA, a second solid support with attached  
CC (differently labelled) monoclonal antibody binding free and bound PSA  
CC (especially selected from (c)) can be used either with or in place of,  
CC the first solid support. The antibodies are useful in cancer screening,  
CC especially prostate and breast cancer. By obtaining total and free PSA  
CC values, their ratio can be used to separate prostatic cancer (Pca) from  
CC benign prostatic hyperplasia (BHP) patients. Measurement of PSA is also  
CC useful after radical prostatectomy, to predict disease persistence. The  
CC antibodies allow evaluation of PSA free/total ratio, enabling separation  
CC of BPH and Pca patients with PSA values 4-10 ng/ml not previously  
CC possible by total PSA testing, avoiding biopsies.  
SQ Sequence 237 AA:

Query Match 100.0%; Score 183; DB 1; Length 237;  
Best Local Similarity 100.0%; Pred. No. 3.20e-13;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 139 EEFLEPKKLCQVDLHVSNVCAOV 163  
QY 1 EEFLEPKKLCQVDLHVSNVCAOV 25

RESULT 5  
ID W03130 standard; Protein; 250 AA.  
AC W03130:  
DT 21-OCT-1996 (first entry)  
DE Prostate-specific antigen.  
KW Prostate-specific antigen; metastasis; carcinoma; tumour; cancer;  
KW neoplasm; archival tissue; polymerase chain reaction; RT-PCR;  
KW primer; kallikrein.  
OS Homo sapiens.  
PN W09621042-A2.  
PD 11-SEP-1995; U00461.  
PF 04-JAN-1996; US-368706.

FT key Location/Qualifiers  
FT peptide 1..13  
FT /label= Pro-peptide  
FT protein 14..250  
FT /label= Mat-protein  
FT /note= "prostate-specific antigen"

PA (UYBO-) UNIV BOSTON.  
 PI Edelman RA, Moreland RB;  
 DR WPI: 96-334017/23.  
 DR N-PSDB: T35867.  
 PT Detecting a metastatic disorder in a fixed biological sample  
 PT comprises reverse transcription PCR of sample nucleic acid and  
 PT detection of metastatic, esp. prostate, tissue-specific target  
 PT sequence  
 PS Disclosure: Fig 1; 60pp; English.  
 CC Prostate-specific antigen (PSA) (W03130) is a serine protease  
 CC (kallikrein) found in both normal and neoplastic prostate  
 CC epithelium. It can be used as a prostate marker. PCR primers  
 CC (see also T35855-62) based on PSA cDNA (T35867) or genomic  
 CC (T35868) sequences can be used for the specific amplification of PSA  
 CC nucleic acids in fresh and archival tissue samples as a means of  
 CC detecting metastatic sequences.  
 SQ Sequence 250 AA;

Query Match 100.0%; Score 183; DB 1; Length 250;  
 Best Local Similarity 100.0%; Pred. No. 3,20e-13;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 152 EEFLLPKKLCQVLDHVISNDVCAOV 176  
 QY 1 EEFLLPKKLCQVLDHVISNDVCAOV 25

RESULT 6  
 ID W13649 standard; Protein: 261 AA.  
 AC W13649;  
 DT 24-FEB-1998 (first entry)  
 DE Human prostatic specific antigen.  
 KW Human; prostatic specific antigen; probasin; rat; cancer;  
 OS Hyperplasia; antisense oligonucleotide.  
 PN WO9711172-A1.  
 PD 27-MAR-1997;  
 PF 20-SEP-1996; U15123.  
 PR 20-SEP-1995; US-004044.  
 PI (MORC-) WORCESTER FOUND BIOMEDICAL RES.  
 PI zamecnik PA;  
 DR WPI: 97-202881/18.  
 DR N-PSDB: T611701.  
 PT Treatment of benign prostatic hyperplasia or prostatic cancer  
 PT using an anti-sense oligonucleotide targeted against prostate  
 PT specific antigen or the probasin gene  
 PS Claim 2; Page 23-30; 49pp; English.  
 CC A novel method has been developed for treating a patient diagnosed as  
 CC having benign prostatic hyperplasia or a prostatic cancer. The method  
 CC comprises administering to the patient a therapeutic amount of a  
 CC composition comprising an antisense oligonucleotide which selectively  
 CC hybridizes to prostate specific antigen (PSA) gene or mRNA or to a  
 CC hybrid gene or mRNA sequence of the patient, where the antisense  
 CC oligonucleotide inhibits expression of the sequence. The present  
 CC sequence represents human PSA. The antisense oligonucleotide can be  
 CC used to effectively inhibit the growth of (and for the killing of)  
 CC hyperplastic cells or cancerous cells of prostatic origin. Expression  
 CC of the PSA and probasin genes is specific to prostatic cells so that the  
 CC antisense oligonucleotide can be administered systemically, making it  
 CC particularly useful in late stage prostatic cancer which has  
 CC metastasized, and in which the cells have become resistant to oestrogen  
 CC or anti-androgen therapy. The method can also be used in benign prostate  
 CC hyperplasia, or early stage prostatic cancer, as an alternative for the  
 CC more radical procedures currently used, such as transurethral resection,  
 CC radical prostatectomy, or physical or chemical castration.  
 SQ Sequence 261 AA;

Query Match 100.0%; Score 183; DB 1; Length 261;  
 Best Local Similarity 100.0%; Pred. No. 3,20e-13;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 163 EEFLLPKKLCQVLDHVISNDVCAOV 187

QY 1 EEFLLPKKLCQVLDHVISNDVCAOV 25

RESULT 7  
 ID R84671 standard; Protein: 237 AA.  
 AC R84671;  
 DT 04-MAR-1996 (first entry)  
 DE Mature kallikrein hK3.  
 KW Kallikrein hK3; serine protease; prostate-specific antigen;  
 KW prostate cancer; diagnosis.  
 OS Homo sapiens.  
 PN WO950758-A1.  
 PD 16-NOV-1995.  
 PF 09-MAY-1995; U06157.  
 PR 10-MAY-1994; US-241174.  
 PR 02-MAY-1995; US-427767.  
 PI (HYBR-) HYBRTECH INC.  
 PI (MAYO-) MAYO FOUNDATION.  
 PI Saeed MS, Tindall DJ, Young CYF;  
 DR WPI: 95-404123/51.  
 PT New isolated prostate-specific kallikrein polypeptide(s) - used to  
 PT develop prods. for use in assays for such polypeptide(s), partic.  
 PT for diagnosis and monitoring of prostate cancer  
 PS Disclosure: Page 31; 61pp; English.  
 CC Human kallikrein hK3 (R84671) is a prostate-specific kallikrein  
 CC showing homology to prostate-specific glandular kallikrein hK2  
 CC (R84667). Nonhomologous regions of the kallikreins can be used  
 CC for prep. of antibodies specific to hK2.  
 SQ Sequence 237 AA;

Query Match 97.3%; Score 178; DB 1; Length 237;  
 Best Local Similarity 96.0%; Pred. No. 1.34e-12;  
 Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DB 139 EEFLLPKKLCQVLDHVISNDVCAOV 163  
 QY 1 EEFLLPKKLCQVLDHVISNDVCAOV 25

RESULT 8  
 ID W45398 standard; Protein: 237 AA.  
 AC W45398;  
 DT 06-JUL-1998 (first entry)  
 DE Prostate-specific antigen protein hK3 (PSA).  
 KW Prostate-specific antigen; hK3; PSA; human; hK2;  
 KW prostate-specific glandular kallikrein; prostate carcinoma;  
 OS Homo sapiens.  
 PN WO9802748-A1.  
 PD 22-JAN-1998.  
 PF 15-JUL-1997; U12322.  
 PR 15-JUL-1996; US-680868.  
 PI (HYBR-) HYBRTECH INC.  
 PI (MAYO-) MAYO FOUNDATION.  
 PI Grauer U, Klee GG, Mikolajczyk SD, Saeed M, Tindall DJ,  
 PI Young CYF;  
 DR WPI: 98-120378/11.  
 PT Diagnostic methods using antibodies which bind prostate antigens  
 PT useful for e.g. monitoring treatment or progression of prostate  
 PT cancer  
 PS Disclosure: Page 55-56; 100pp; English.  
 CC This polypeptide comprises human prostate-specific antigen (PSA)  
 CC hK3. PSA levels are widely used as a prognostic indicator of  
 CC prostate carcinoma. The invention provides a novel diagnostic  
 CC method comprising contacting antibodies that specifically bind to  
 CC human prostate-specific glandular kallikrein protein hK2 (see  
 CC W45395) or pro-hK2 (hK2, see W45396), but not with PSA, with a  
 CC sample of physiological fluid from a human. The assay is based on  
 CC the discovery that hK2 is detected in the supernatant of a  
 CC prostate carcinoma cell line and that hK2 is present in human  
 CC physiological fluid from prostate cancer cells. The method is  
 CC useful for monitoring the treatment and/or progression of prostate  
 CC cancer, or for the early detection of prostate cancer in males that

CC have BPH or a high grade prostatic neoplasia (HPGN) or whose family  
CC members have or had BPH, HPGN or a prostate cancer. Monitoring the  
CC presence and/or amount of hK2 complexes with plasma proteins may BPH.  
CC also be important in distinguishing between prostate cancer and BPH.  
SQ Sequence 237 AA;

Query Match 97.3%; Score 178; DB 1; Length 237;  
Best Local Similarity 96.0%; Pred. No. 1.34e-12;  
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 139 EEFLEPKKLCVQLHVISNDVCAOV 163  
1 EEFLEPKKLCVQLHVISNDVCAOV 25

RESULT 9  
ID W27387 standard; peptide: 30 AA.

AC W27387;  
DE 14-APR-1998 (first entry)  
DE Prostate specific antigen oligo-epitope peptide.  
KW Prostate specific antigen; PSA; epitope; vaccine; immune response;  
KW Prostate tumor cell; human leukocyte antigen; HLA; cytotoxic;  
KW T lymphocyte; cancer.  
OS Synthetic.  
OS Homo sapiens.  
PN MO9735021-AA.  
PD 25-SEP-1997.  
PE 19-MAR-1996; US-618936.  
PR 20-MAR-1996; US-618936.  
PA (USSH) US SEC DEPT HEALTH.  
PI Schlom J, Tsang K, Zaremba S;  
DR MPI: 97-480226/44.  
DR N-PSDB: T90073.  
PT Prostate specific antigen multiple epitope peptide - useful in  
PT vaccine to produce immune response against prostate specific antigen  
PT in prostate tumor cell  
PS Claim 13; Page 63; 85pp; English.  
CC The present sequence represents a prostate specific antigen  
CC oligo-epitope peptide (PSA-OP). The PSA-OP comprises more than 1  
CC adjoined PSA-OP which generates a prostate specific response in  
CC a portion of the human population. The PSA-OP can be used to produce  
CC PSA specific cytotoxic T lymphocytes, in vivo or in vitro, which can  
CC kill or prevent the establishment or growth of PSA positive tumor  
CC cells, i.e. to prevent or treat prostatic cancer. The products may be  
CC used to elicit a PSA specific immune response. When a recombinant virus  
CC is used to elicit the response it is coadministered with a carrier, a  
CC biological response modifier e.g. Interleukin 2, 6 or 12, interferon,  
CC tumor necrosis factor, granulocyte/macrophage-colony stimulating factor  
CC or cyclophosphamide and an adjuvant selected from RIBI Detox, QS21,  
CC alum and incomplete Freund's adjuvant. PSA-OP comprises multiple PSA  
CC epitopes, allowing an immune response to PSA in a population of humans  
CC having more than 1 HLA class I molecule type to be generated.  
SQ Sequence 30 AA;

8/3

Query Match 93.4%; Score 171; DB 1; Length 30;  
Best Local Similarity 100.0%; Pred. No. 9.87e-12;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 FLTPKRLQCVDLHVISNDVCAOV 23  
3 FLTPKRLQCVDLHVISNDVCAOV 25

RESULT 10  
ID W11023 standard; Protein: 240 AA.

AC W11023;  
DE 27-OCT-1997 (first entry)  
DE Human prostate specific antigen.  
KW PSA; kallikrein; prostate cancer; antigenic; polyclonal antisera.  
OS Homo sapiens.  
PN MO9640754-AA.  
PD 19-DEC-1996.

PF 06-JUN-1996; U09303.  
PR 07-JUN-1995; US-472228.  
PA (TEXA) UNIV TEXAS SYSTEM.  
PI Fritsche HA, Johnston DA, Kokolus WJ;  
DR MPI: 97-108633/10.  
PT Antigen peptide derived from prostate specific antigen - does not  
PT cross react with related kallikreins, for diagnosis of prostate  
PT cancer.  
PS Claim 1; Fig 1; 74pp; English.  
CC The present sequence represents human prostate specific antigen (PSA)  
CC which is used to produce novel peptides derived from the 240 residue  
CC sequence. The preferred peptides have two hydrophobic regions and one  
CC hydrophilic region each of about 5 aa in length, arranged as follows:  
CC hydrophobic-hydrophilic-hydrophobic. Polyclonal antisera specific for  
CC the antigenic peptides may be used in a method for diagnosing prostate  
CC cancer in vivo or in vitro. The peptides represent a sequence unique  
CC to PSA which does not cross react with certain kallikreins.  
SQ Sequence 240 AA;

Query Match 73.2%; Score 134; DB 1; Length 240;  
Best Local Similarity 94.7%; Pred. No. 3.20e-07;  
Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 148 KKLQCVQLHVISNDVCAOV 166  
7 KKLQCVQLHVISNDVCAOV 25

RESULT 11  
ID W96186 standard; Protein: 237 AA.  
AC W96186;  
DE 27-APR-1999 (first entry)  
DE Mature human kallikrein 2 (hK2).  
KW Human kallikrein 2; hK2; breast cancer; diagnosis; detection;  
KW treatment; monitoring; prostate specific antigen; PSA.  
OS Homo sapiens.  
PN WO9859073-A1.  
PD 30-DEC-1998.  
PE 19-JUN-1998; U12840.  
PR 20-JUN-1997; US-050963.  
PA (KLEE) KLEE G G.  
PA (MAYO) MAYO FOUNDATION.  
PA (TIND) TINDALL D J.  
PA (YOUNG) YOUNG C Y F.  
PI Klee GG, Tindall DJ, Young CF;  
DR MPI: 99-105632/09.  
DR N-PSDB: X08946.

PT Use of human kallikrein 2 - as a marker for developing products for  
PT the diagnosis, prognosis, monitoring and treatment of breast cancer  
PT disclosure; Page 31; 67pp; English.  
PS Human kallikrein 2 (hK2) is expressed at elevated levels relative to  
PS the prostate cancer antigen, prostate specific antigen (PSA) by  
PS breast cancer cells when stimulated by androgens. Detecting levels  
PS of hK2 may provide a simple diagnostic tool for detecting or  
PS determining breast cancer. Detecting hK2 is achieved by producing  
PS an hK2 DNA from hK2 RNA by reverse transcription. The hK2 DNA  
PS obtained is then amplified by PCR and detected using probes.  
PS Determination of breast cancer may also be determined by exposing  
PS the hK2 polypeptide to an agent which binds to it and then detecting  
PS the binary complex formed. The amount of complex formed correlates  
PS directly to the presence of breast cancer in the human individual.  
PS The methods can be used in the diagnosis, treatment and/or  
PS monitoring of the progression or remission of breast cancer and/or  
PS monitoring hK2 levels.  
SQ Sequence 237 AA;

Query Match 71.0%; Score 130; DB 1; Length 237;  
Best Local Similarity 66.7%; Pred. No. 9.62e-07;  
Matches 16; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Db 139 EEFLEPKKLCVQLHVISNDVCAOV 162  
1 EEFLEPKKLCVQLHVISNDVCAOV 24





KM prostate cancer: detection: hk2; hk1; hk3; phk2; tissue kallikrein;  
 KW pphk2; prostate-specific glandular kallikrein protein; PSA; human;  
 KW prostate-specific antigen.  
 OS Homo sapiens.  
 OS Synthetic.  
 PN MO9846795-A1.  
 PD 22-OCT-1998.  
 PF 09-APR-1998: 007027.  
 PR 11-APR-1997; US-843076.  
 PA (BAYO ) BAYLOR COLLEGE MEDICINE.  
 PA (MAYO-) MAYO FOUNDATION.  
 PI Slawin KM, Tindall DJ, Young CYF;  
 DR WPI; 98-594592/50.  
 PT Detection of human kallikrein 2 RNA - by reverse transcription and  
 PT amplification by PCR, for detecting, monitoring and staging of  
 PT prostate cancer.  
 PS Disclosure: Page 70-71; 90pp; English.  
 CC The present invention describes a diagnostic method for detecting human  
 CC kallikrein 2 (hk2) DNA. The method comprises: (a) contacting DNA obtained  
 CC by reverse transcription (RT) of RNA from a human physiological sample  
 CC which comprises cells suspected of containing hk2 RNA with at least 2  
 CC oligonucleotides to amplify the DNA by PCR to yield amplified hk2 DNA,  
 CC where the conditions amplify the DNA obtained by RT of RNA from at least  
 CC one cell containing hk2 in a sample which comprises at least 107 to 109  
 CC cells; and (b) detecting the presence of the amplified hk2 DNA. The  
 CC method can be used for detecting, monitoring the progression of and  
 CC pathologically staging prostate cancer. The present sequence represents  
 CC a hk2 variant.  
 SQ Sequence 237 AA;

Query Match 71.0%; Score 130; DB 1; Length 237;  
 Best Local Similarity 66.7%; Pred. No. 9.62e-07;  
 Matches 16; Conservative 5; Mismatches 3; Indels 0; Gaps 0;  
 Db 139 EEFLRPSIOCVSLHLISNDWCAR 162  
 QY 1 EEFLPKRLQCVDLHLVISNDVCAQ 24

Search completed: Thu Jun 22 21:23:30 2000  
 Job time : 8 secs.

(W.T.)

Distribution rights by Oxford Molecular Ltd

in - protein database search,

2000, 4:57:47

generated.

US-09-332-866-1

EEFLT PKKLQCVDLHVISNDVCAOV 25

AM 150

1443/480 residues

### Listing first 45 summaries

## Issue

variance 64.4332; scale 0.307

### Analysis of the local score distribution

## SUMMARIES

length	ID	Description	pred. No
237	US-08-844	Sequence 2, Applicatio	3.60e-11
237	US-08-718	Sequence 2, Applicatio	3.60e-11
237	US-08-768	Sequence 1, Applicatio	3.60e-11
261	US-08-744	Sequence 5, Applicatio	3.60e-11
261	US-08-102	Sequence 5, Applicatio	3.60e-11
237	PCT-US94-0	Sequence 11, Applicati	1.45e-12
237	US-08-096	Sequence 11, Applicati	1.45e-12
237	PCT-US95-0	Sequence 11, Applicatio	1.45e-12
240	US-08-760	Sequence 1, Applicatio	2.58e-07
240	PCT-US95-0	Sequence 1, Applicatio	2.58e-07
237	US-08-472	Sequence 1, Applicatio	2.58e-07
237	US-08-056	Sequence 10, Applicati	7.58e-07
237	PCT-US94-0	Sequence 10, Applicati	7.58e-07
237	US-08-768	Sequence 16, Applicati	7.58e-07
237	US-08-768	Sequence 16, Applicati	7.58e-07
237	US-08-768	Sequence 21, Applicati	7.58e-07
238	PCT-US95-0	Sequence 8, Applicatio	7.58e-07
238	US-08-768	Sequence 8, Applicatio	7.58e-07
244	PCT-US95-0	Sequence 10, Applicati	7.58e-07
244	US-08-768	Sequence 10, Applicati	7.58e-07
251	PCT-US95-0	Sequence 6, Applicatio	7.58e-07
251	US-08-768	Sequence 6, Applicatio	7.58e-07
251	US-08-768	Sequence 19, Applicatio	7.58e-07
18	PCT-US94-0	Sequence 4, Applicatio	3.47e-03

1	US-08-096-	Sequence 4,	Applicatio	3.47e-01
1	US-08-472-	Sequence 22,	Applicati	5.76e-03
1	PCR-US96-0	Sequence 22,	Applicati	5.76e-03
1	US-08-744-	Sequence 3,	Applicatio	7.47e-03
2	US-09-103-	Sequence 3,	Applicatio	7.47e-03
2	US-08-824-	Sequence 3,	Applicatio	2.07e-02
2	US-08-681-	Sequence 4,	Applicatio	2.07e-02
2	US-08-790-	Sequence 4,	Applicatio	2.07e-02
2	US-09-102-	Sequence 3,	Applicatio	2.07e-02
3	US-08-744-	Sequence 4,	Applicatio	2.07e-02
3	US-08-807-	Sequence 4,	Applicatio	2.07e-02
2	US-08-750-	Sequence 4,	Applicatio	2.07e-02
2	US-08-834-	Sequence 1,	Applicatio	2.07e-02
3	US-08-807-	Sequence 5,	Applicatio	2.07e-02
2	US-08-750-	Sequence 4,	Applicatio	2.07e-02
2	US-08-897-	Sequence 31,	Applicati	7.28e-02
1	US-09-102-	Sequence 1,	Applicatio	7.28e-02
1	US-08-744-	Sequence 1,	Applicatio	7.28e-02
1	US-08-472-	Sequence 7,	Applicatio	1.97e-01
1	US-08-615-	Sequence 7,	Applicatio	1.97e-01
1	US-08-218-	Sequence 7,	Applicatio	1.97e-01
3	US-09-074-	Sequence 7,	Applicatio	1.97e-01

## ALIGNMENTS

PRJ: 237 AA

05/08844024

03/06644024

ron E., et al

Prostate Cancer

555

Miller Plaza

PC-DOS/MS-DOS

DATA: 11/11/11

5

08/229,391

**P.**

NUMBER: 43677/JPW/TEP

0525

NO: 2

ENTER

CC TOPOLOGY: linear  
CC MOLECULE TYPE: Protein  
SQ SEQUENCE 237 AA; 26151 MW; 312585 CN;  
Query Match 100.0%; Score 183; DB 2; Length 237;  
Best Local Similarity 100.0%; Pred. No. 3,60e-13;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 139 EEFLLPKKLCQVDLHVISNDVCAQV 163  
1 EEFLLPKKLCQVDLHVISNDVCAQV 25  
OY 1 EEFLLPKKLCQVDLHVISNDVCAQV 25  
RESULT 2  
ID US-08-718-547-2 STANDARD: PRT; 237 AA.  
AC xxxxxx  
DE Sequence 2, Application US/08718547  
CC Sequence 2, Application US/08718547  
CC Patent No. 5976794  
CC GENERAL INFORMATION:  
CC APPLICANT: Katz, Aaron E.; Buttyan, Ralph; Ratfo, Anthony;  
CC TITLE OF INVENTION: A Method For Molecular Staging Of  
CC NUMBER OF SEQUENCES: 4  
CC CORRESPONDENCE ADDRESS:  
CC ATTORNEY/AGENT INFORMATION:  
CC STREET: 1185 Avenue of the Americas  
CC CITY: New York  
CC STATE: New York  
CC COUNTRY: U.S.A.  
CC ZIP: 10036  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patentin Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/718,547  
CC FILING DATE:  
CC CLASSIFICATION: 435  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: White, John P.  
CC REGISTRATION NUMBER: 28,678  
CC REFERENCE/DOCKET NUMBER: 43677-A-PCT/JPW/MUG  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (212) 278-0400  
CC TELEFAX: (212) 391-0525  
CC TELEX:  
CC INFORMATION FOR SEQ ID NO: 2:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 237 amino acids  
CC TYPE: amino acid  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
SQ SEQUENCE 237 AA; 26151 MW; 312585 CN;  
Query Match 100.0%; Score 183; DB 2; Length 237;  
Best Local Similarity 100.0%; Pred. No. 3,60e-13;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 139 EEFLLPKKLCQVDLHVISNDVCAQV 163  
1 EEFLLPKKLCQVDLHVISNDVCAQV 25  
OY 1 EEFLLPKKLCQVDLHVISNDVCAQV 25  
RESULT 3  
ID US-08-768-859A-1 STANDARD: PRT; 237 AA.

XX  
AC xxxxxx  
AC  
DT  
DE Sequence 1, Application US/08768859A  
XX Sequence 1, Application US/08768859A  
XX Patent No. 6013471  
CC GENERAL INFORMATION:  
CC APPLICANT: Tindall, Donald J.  
CC APPLICANT: Young, Charles Y.F.  
CC APPLICANT: Saeedi, Mohammed S.  
CC TITLE OF INVENTION: DNA Encoding HK2 Variant Polypeptides  
CC NUMBER OF SEQUENCES: 21  
CC CORRESPONDENCE ADDRESS:  
CC ATTORNEY/AGENT INFORMATION:  
CC STREET: P.O. Box 2938  
CC CITY: Minneapolis  
CC STATE: MN  
CC COUNTRY: USA  
CC ZIP: 55402  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patentin Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/768,859A  
CC FILING DATE: 17-DECEMBER-1995  
CC CLASSIFICATION: 435  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Woessner, Warren D.  
CC REGISTRATION NUMBER: 30,440  
CC REFERENCE/DOCKET NUMBER: 545,002US3  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 612-339-0361  
CC TELEFAX: 612-339-0361  
CC INFORMATION FOR SEQ ID NO: 1:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 237 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: peptide  
SQ SEQUENCE 237 AA; 26089 MW; 312729 CN;  
Query Match 100.0%; Score 183; DB 3; Length 237;  
Best Local Similarity 100.0%; Pred. No. 3,60e-13;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 139 EEFLLPKKLCQVDLHVISNDVCAQV 163  
1 EEFLLPKKLCQVDLHVISNDVCAQV 25  
OY 1 EEFLLPKKLCQVDLHVISNDVCAQV 25  
RESULT 4  
ID US-08-744-026-5 STANDARD: PRT; 261 AA.  
AC xxxxxx  
AC  
DT  
DE Sequence 5, Application US/08744026  
XX Sequence 5, Application US/08744026  
XX Patent No. 5786148  
CC GENERAL INFORMATION:  
CC APPLICANT: Bandman, Olga  
CC APPLICANT: Goli, Surya K.  
CC TITLE OF INVENTION: A NOVEL PROSTATE-SPECIFIC  
CC NUMBER OF SEQUENCES: 5  
CC

```
CC CC CORRESPONDENCE ADDRESS:  
CC CC ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
CC CC STREET: 3174 Porter Drive  
CC CC CITY: Palo Alto  
CC CC STATE: CA  
CC CC COUNTRY: US  
CC CC ZIP: 94304  
CC CC COMPUTER READABLE FORM:  
CC CC MEDIUM TYPE: Diskette  
CC CC OPERATING SYSTEM: DOS  
CC CC SOFTWARE: FastSeq Version 2.0  
CC CC CURRENT APPLICATION DATA:  
CC CC FILING DATE: US/08/744, 026  
CC CC CLASSIFICATION: 514  
CC CC PRIOR APPLICATION DATA:  
CC CC APPLICATION NUMBER:  
CC CC FILING DATE:  
CC CC ATTORNEY/AGENT INFORMATION:  
CC CC NAME: Billings, Lucy J  
CC CC REGISTRATION NUMBER: 36,749  
CC CC REFERENCE/DOCKET NUMBER: PF-0154 US  
CC CC TELECOMMUNICATION INFORMATION:  
CC CC TELEPHONE: 415-855-0555  
CC CC TELEFAX: 415-845-4166  
CC CC INFORMATION FOR SEQ ID NO: 5:  
CC CC SEQUENCE CHARACTERISTICS:  
CC CC LENGTH: 261 amino acids  
CC CC TYPE: amino acid  
CC CC STRANDEDNESS: single  
CC CC TOPOLOGY: linear  
CC CC IMMEDIATE SOURCE:  
CC CC LIBRARY: GenBank  
CC CC CLONE: 190553  
CC CC SEQUENCE 261 AA; 28741 MW; 378398 CN;  
  
Query Match 100.0%; Score 183; DB 1; Length 261;  
Best Local Similarity 100.0%; Pred.No. 3,60e-13;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 163 EEFITPKKLOCVDLHVISNDVCAQV 187  
QY 1 EEFITPKKLOCVDLHVISNDVCAQV 25  
|||||  
  
RESULT 5  
ID US-09-102-732-5 STANDARD: PRT; 261 AA.  
AC xxxxxx  
DT  
XX  
XX  
DE Sequence 5, Application US/09102732  
Sequence 5, Application US/09102732  
Patent No. 5922321  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
APPLICANT: Goll, Susya K.  
TITLE OF INVENTION: A NOVEL PROSTATE-SPECIFIC  
TITLE OF INVENTION: KALIKREIN  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSER: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: US  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible
```

```

CC      OPERATING SYSTEM: DOS
CC      SOFTWARE: FastSD version 2.0
CC      CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER: US/09/102,732
CC      FILING DATE:
CC      CLASSIFICATION:
CC      PRIOR APPLICATION DATA:
CC      APPLICATION NUMBER: 08/744,036
CC      FILING DATE:
CC      ATTORNEY/AGENT INFORMATION:
CC      NAME: Billings, Lucy J
CC      REGISTRATION NUMBER: 36,749
CC      REFERENCE/DOCKET NUMBER: PF-0154 US
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE: 415-855-0555
CC      TELEFAX: 415-845-4166
CC      INFORMATION FOR SEQ ID NO: 5:
CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH: 261 amino acids
CC      TYPE: amino acid
CC      STRANDEDNESS: single
CC      TOPOLOGY: linear
CC      IMMEDIATE SOURCE:
CC      LIBRARY: GenBank
CC      CLONE: 190553
CC      SEQUENCE 261 AA; 28741 MW; 378398 CN;

Query Match      100.0%; Score 183; DB 2; Length 261;
Best Local Similarity 100.0%; Pred. No. 3,60e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      163 EEF1PKKLCQVCDLHVISNDVCAQV 187
Qy      1 EEF1PKKLCQVCDLHVISNDVCAQV 25
      |||||||||||||||||||
RESULT 6          STANDARD:      PRT: 237 AA.
ID      PCT-US94-07329-11
XX      xxxxxx
XX

Sequence 11, Application PC/TUS9407329

Sequence 11, Application PC/TUS9407329
GENERAL INFORMATION:
APPLICANT: Mayo Foundation for Medical
APPLICANT: Education and Research
TITLE OF INVENTION: Antibodies Specific for Human
TITLE OF INVENTION: Prostate Glandular Kallikrein
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSER: Schwegman, Lundberg & Woessner
STREET: 3500 IDS Center
STREET: 80 South Eighth Street
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55402

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/07329
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Raasch, Kevin W.
REGISTRATION NUMBER: 35,651
REFERENCE/DOCKET NUMBER: 150,63WO

```



```

XX AC      XX
XX XX      XXXXX
XX DT
XX XX
XX DE      Sequence 1, Application PC/TUS9609303
XX XX      CC
CC CC      Sequence 1, Application PC/TUS9609303
CC CC      GENERAL INFORMATION:
CC CC      APPLICANT:
CC CC      TITLE OF INVENTION: IMMUNOGENIC PEPTIDES OF PROSTATE
CC CC      TITLE OF INVENTION: SPECIFIC ANTIGEN
CC CC      NUMBER OF SEQUENCES: 22
CC CC      COMPUTER READABLE FORM:
CC CC      MEDIUM TYPE: Floppy disk
CC CC      OPERATING SYSTEM: PC-DOS/MS-DOS
CC CC      SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CC CC      CURRENT APPLICATION DATA:
CC CC      APPLICATION NUMBER: PCT/US96/09303
CC CC      FILING DATE:
CC CC      CLASSIFICATION:
CC CC      PRIOR APPLICATION DATA:
CC CC      APPLICATION NUMBER: US 08/472,228
CC CC      FILING DATE: 07-JUN-1995
CC CC      INFORMATION FOR SEQ ID NO: 1:
CC CC      SEQUENCE CHARACTERISTICS:
CC CC      LENGTH: 240 amino acids
CC CC      TYPE: amino acid
CC CC      STRANDEDNESS:
CC CC      TOPOLOGY: linear
CC CC      SEQUENCE 240 AA: 26550 MW: 319046 CN;

Query Match          73.2%  Score 134; DB 4; Length 240;
Best Local Similarity 94.7%; Pred. No. 2.58e-07;
Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db      148 KRLQCVQLHVVISNDVCAQY 166
        |||||:|||||
        7 KRLQCVDLHVISNDVCAQY 25

RESULT 10
ID      US-08-472-228A-1          STANDARD:          PRT:      240 AA.
XX      XXXXXX

Sequence 1, Application US/08472228A

Sequence 1, Application US/08472228A
Patent No. 5807978
GENERAL INFORMATION:
APPLICANT: Kokolus, William J.
APPLICANT: Fritzsche, Herbert A.
APPLICANT: Johnston, Dennis A.
TITLE OF INVENTION: IMMUNOGENIC PEPTIDES OF PROSTATE
TITLE OF INVENTION: SPECIFIC ANTIGEN
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESSES:
ADDRESSER: Arnold, White & Durkee
STREET: P.O. Box 4433
City: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

```

```
CC CC APPLICATION NUMBER: US/08/472,228A
CC CC FILING DATE: 07-JUN-1995
CC CLASSIFICATION: 424
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Highlander, Steven L.
CC REGISTRATION NUMBER: 37,642
CC REFERENCE/DOCKET NUMBER: UTSC:388
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (512) 418-3000
CC TELEFAX: (512) 474-7577
CC TEXT: 79-0924
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 240 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 240 AA; 26550 MW; 319046 CN;
CC
CC Query Match 73.2%; Score 134; DB 1; Length 240;
CC Best Local Similarity 94.7%; Pred. No. 2,58e+07;
CC Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
CC
CC Db 148 KKLQCVOLHVISNDVCAOV 166
CC |||||:|||||
CC 7 KKLQCVDLHVISNDVCAOV 25
CC
CC RESULT 11
CC ID US-08-096-946-10 STANDARD; PRT; 237 AA.
CC XX xxxxxx
CC XX
CC DE Sequence 10, Application US/08096946
CC
CC Sequence 10, Application US/08096946
CC Patent No. 5516639
CC GENERAL INFORMATION:
CC APPLICANT: Tindall, Donald J.
CC APPLICANT: Young, Charles Y-F
CC APPLICANT: Klee, George G.
CC TITLE OF INVENTION: Antibodies Specific for Human
CC NUMBER OF INVENTIONS: Prostate Glandular kallikrein
CC NUMBER OF SEQUENCES: 11
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Merchant & Gould
CC STREET: 3100 No. 5516639west Center
CC CITY: Minneapolis
CC STATE: MN
CC COUNTRY: USA
CC ZIP: 55402
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/096,946
CC FILING DATE: 19930722
CC CLASSIFICATION: 436
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Woessner, Warren D.
CC REGISTRATION NUMBER: 30,440
CC REFERENCE/DOCKET NUMBER: 1016.62-US-01
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 612-332-5300
CC TELEFAX: 612-332-8081
CC INFORMATION FOR SEQ ID NO: 10:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 237 amino acids
```



CC TYPE: amino acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: peptide  
CC IMMEDIATE SOURCE:  
CC CLONE: Amino acid sequence of hK2  
CC CLONE: (deduced from cDNA sequence)  
SQ SEQUENCE 237 AA; 26182 MW; 304329 CN;

Query Match 71.0%; Score 130; DB 1; Length 237;  
Best Local Similarity 66.7%; Pred. No. 7.58e-07;  
Matches 16; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Db 139 EEFLEPRSLQCVSLHLNSDMCAR 162  
1 EEFLEPRSLQCVSLHLNSDMCAR 24

RESULT 12  
ID PCT-US94-07329-10 STANDARD: PRT: 237 AA.  
XX xxxxxx

Sequence 10, Application PC/TUS9407329

CC GENERAL INFORMATION:  
CC APPLICANT: Mayo Foundation for Medical  
CC APPLICANT: Education and Research  
CC TIME OF INVENTION: Antibodies Specific for Human  
CC TITLE OF INVENTION: Prostate Glandular kallikrein  
CC NUMBER OF SEQUENCES: 11  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Schwegman, Lundberg & Woessner  
CC STREET: 3500 IDS Center  
CC STREET: 80 South Eighth Street  
CC CITY: Minneapolis  
CC STATE: MN  
CC COUNTRY: USA  
CC ZIP: 55402  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: PCT/US94/07329

CC FILING DATE:  
CC CLASSIFICATION:  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Raasch, Kevin W.  
CC REGISTRATION NUMBER: 35,651  
CC REFERENCE/DOCKET NUMBER: 150.62WO  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 612-339-0331  
CC INFORMATION FOR SEQ ID NO: 10:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 237 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: peptide  
CC IMMEDIATE SOURCE:  
CC CLONE: Amino acid sequence of hK2  
CC CLONE: (deduced from cDNA sequence)

SQ SEQUENCE 237 AA; 26182 MW; 304329 CN;

Query Match 71.0%; Score 130; DB 4; Length 237;  
Best Local Similarity 66.7%; Pred. No. 7.58e-07;  
Matches 16; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Db 139 EEFLEPRSLQCVSLHLNSDMCAR 162  
1 EEFLEPRSLQCVSLHLNSDMCAR 24

RESULT 13  
ID PCT-US95-06157-16 STANDARD: PRT: 237 AA.  
XX xxxxxx

Sequence 16, Application PC/TUS9506157

CC GENERAL INFORMATION:  
CC APPLICANT: Mayo Foundation for Medical Education  
CC APPLICANT: and Research  
CC APPLICANT: Hydritech Incorporated  
CC APPLICANT: Tindall, Donald J.  
CC APPLICANT: Young, Charles Y.F.  
CC APPLICANT: Saeedi, Mohammed S.  
CC TITLE OF INVENTION: Recombinant hK2 Polypeptide  
CC NUMBER OF SEQUENCES: 18  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Schwegman, Lundberg & Woessner, P.A.  
CC STREET: 3500 IDS Center  
CC CITY: Minneapolis  
CC STATE: MN  
CC COUNTRY: USA  
CC ZIP: 55402  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: PCT/US95/06157

CC FILING DATE:  
CC CLASSIFICATION:  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Raasch, Kevin W.  
CC REGISTRATION NUMBER: 35,561  
CC REFERENCE/DOCKET NUMBER: 150.148WO1  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 612-339-0331  
CC INFORMATION FOR SEQ ID NO: 16:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 237 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: peptide  
CC SEQUENCE 237 AA; 26182 MW; 304329 CN;

Query Match 71.0%; Score 130; DB 4; Length 237;  
Best Local Similarity 66.7%; Pred. No. 7.58e-07;  
Matches 16; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Db 139 EEFLEPRSLQCVSLHLNSDMCAR 162  
1 EEFLEPRSLQCVSLHLNSDMCAR 24

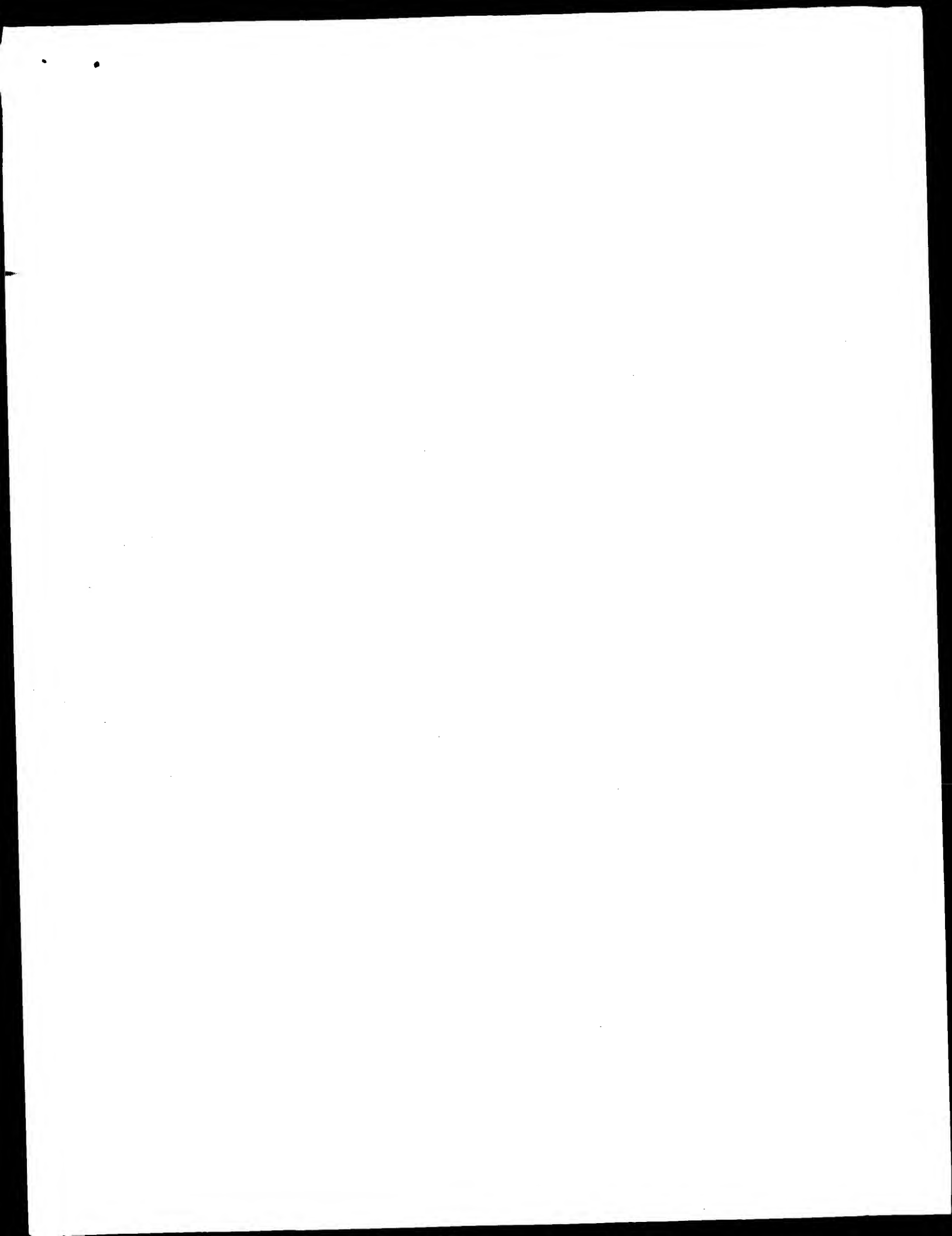
RESULT 14  
ID US-08-768-859A-21 STANDARD: PRT: 237 AA.  
XX xxxxxx

DT

DE Sequence 21, Application US/08768859A  
XX  
CC Sequence 21, Application US/08768859A  
CC Patent No. 6013471  
CC GENERAL INFORMATION:  
CC APPLICANT: Tindall, Donald J.  
CC APPLICANT: Young, Charles Y.F.  
CC APPLICANT: Saeed, Mohammed S.  
CC TITLE OF INVENTION: DNA Encoding HK2 Variant Polypeptides  
CC NUMBER OF SEQUENCES: 21  
CC CORRESPONDENCE ADDRESSES:  
CC ADDRESS: Schwegman, Lundberg, Woessner & Kluth, P.A.  
CC STREET: P.O. Box 2938  
CC CITY: Minneapolis  
CC STATE: MN  
CC COUNTRY: USA  
CC ZIP: 55402  
CC  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/768,859A  
CC FILING DATE: 17-DECEMBER-1995  
CC CLASSIFICATION: 435  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Woessner, Warren D.  
CC REGISTRATION NUMBER: 30,440  
CC REFERENCE/DOCKET NUMBER: 545.002US3  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 612-339-0331  
CC TELEFAX: 612-339-3061  
CC INFORMATION FOR SEQ ID NO: 21:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 237 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: peptide  
CC SEQUENCE 237 AA: 26187 MW: 308471 CN;  
SO  
Query Match 71.0%; Score 130; DB 3; Length 237;  
Best Local Similarity 66.7%; Pred. No. 7.58e-07;  
Matches 16; Conservative 5; Mismatches 3; Indels 0; Gaps 0;  
DB 139 EEFLLPRSLQCVSLHLSNDKCAR 162  
QY 1 EEFLLTPKRLQCVDLHVISNDVCAQ 24  
RESULT 15  
ID US-08-768-859A-16 STANDARD: PRT: 237 AA.  
AC xxxxxx  
DT  
XX  
DE Sequence 16, Application US/08768859A  
XX  
CC Sequence 16, Application US/08768859A  
CC Patent No. 6013471  
CC GENERAL INFORMATION:  
CC APPLICANT: Tindall, Donald J.  
CC APPLICANT: Young, Charles Y.F.  
CC APPLICANT: Saeed, Mohammed S.  
CC TITLE OF INVENTION: DNA Encoding HK2 Variant Polypeptides  
CC NUMBER OF SEQUENCES: 21  
CC CORRESPONDENCE ADDRESSES:  
CC ADDRESS: Schwegman, Lundberg, Woessner & Kluth, P.A.  
CC STREET: P.O. Box 2938  
CC CITY: Minneapolis  
CC STATE: MN

CC COUNTRY: USA  
CC ZIP: 55402  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/768,859A  
CC FILING DATE: 17-DECEMBER-1995  
CC CLASSIFICATION: 435  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Woessner, Warren D.  
CC REGISTRATION NUMBER: 30,440  
CC REFERENCE/DOCKET NUMBER: 545.002US3  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 612-339-0331  
CC TELEFAX: 612-339-3061  
CC INFORMATION FOR SEQ ID NO: 16:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 237 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: peptide  
CC SEQUENCE 237 AA: 26159 MW: 304348 CN;  
SO  
Query Match 71.0%; Score 130; DB 3; Length 237;  
Best Local Similarity 66.7%; Pred. No. 7.58e-07;  
Matches 16; Conservative 5; Mismatches 3; Indels 0; Gaps 0;  
DB 139 EEFLLPRSLQCVSLHLSNDKCAR 162  
QY 1 EEFLLTPKRLQCVDLHVISNDVCAQ 24

Search completed: Thu Jun 22 21:23:55 2000  
Job time : 8 secs.



TRANSLATION AA-&gt;NA

(TM)

Release 3.1A John F. Collins, Bioinformatics Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

Mparch\_tpm n.a. - n.a. Smith-Waterman search, using a protein query  
which has been backtranslated into n.a. using IUPAC symbols

Run on: Fri Jun 23 15:48:17 2000; Maspar time 285.75 seconds

Tabular output not generated. 1118.717 Million cell updates/sec

Title: >US-09-332-866-1  
Description: (123) from US09332866.pep  
Perfect Score: 316  
N.A. Sequence: 1 GARGARTTYVTNACNCNNA.....AYGAYTNGYGCNCARGTN 75  
Comp: CYYCTYAAARANTWGNGNTT.....TTCCTCANACACGNGTTCAN

Scoring table: TABLE bkttranslated2  
Gap 40

Match STD : Dbase 0; Query 0

Searched: 23132 seqs, 213114489 bases x 2

Post-processing: Minimum Match 08

Listing first 45 summaries

Database:

emb161  
1:em\_ba1 2:em\_ba2 3:em\_fun 4:em\_htg1 5:em\_htg2 6:em\_htg3  
7:em\_hum1 8:em\_hum2 9:em\_hum3 10:em\_hum4 11:em\_in  
12:em\_com 13:em\_ov 14:em\_ov 15:em\_pat 16:em\_pi 17:em\_ro  
18:em\_sts 19:em\_un 20:em\_v1  
genbank16  
21:gb\_htg1 22:gb\_htg10 23:gb\_htg11 24:gb\_htg12  
25:gb\_htg13 26:gb\_htg14 27:gb\_htg15 28:gb\_htg16 29:gb\_htg17  
30:gb\_htg18 31:gb\_htg19 32:gb\_htg20 33:gb\_htg21 34:gb\_htg22

Database:

Statistics: Mean 78.372; Variance 327.890; scale 0.239

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result	No.	Score	Query	Match	Length	DB	ID	Description	Pred. No.
	1	277	87.7	40334	28	AC011523	Homo sapiens chromosome	7.33e-08	
	2	145	45.9	40871	31	AC020019	Drosophila melanogaste	5.45e+01	
	3	137	43.4	2286	31	AC019591	Drosophila melanogaste	1.43e+02	
	4	137	43.4	60074	29	AC013501	Homo sapiens chromosome	1.43e+02	
C	5	137	43.4	174801	29	AC014376	Drosophila melanogaste	1.43e+02	
	6	137	43.4	200319	27	AL135931	Homo sapiens chromosome	1.43e+02	
	7	137	43.4	228527	22	AC013342	Homo sapiens clone Rpl	1.43e+02	
C	8	135	42.7	135499	24	AC008969	Homo sapiens chromosome	1.80e+02	
	9	132	41.8	17438	29	AC014821	Drosophila melanogaste	2.53e+02	
C	10	132	41.8	148162	33	AC012865	Homo sapiens clone Rpl	2.53e+02	
	11	132	41.8	149408	31	AC013582	Homo sapiens clone Rpl	2.53e+02	
C	12	132	41.8	151670	28	AC007777	Homo sapiens chromosome	2.53e+02	

13	132	41.8	156492	32	AC008460	Homo sapiens chromosome 2	5.3e+02
14	132	41.8	182606	27	AL136986	Homo sapiens chromosome 2	5.3e+02
15	132	41.8	202335	33	AC020603	Homo sapiens chromosome 2	5.3e+02
16	131	41.5	48836	31	AC020151	Drosophila melanogaster	2.83e+02
17	131	41.5	112958	29	AC010107	Drosophila melanogaster	2.83e+02
18	131	41.5	129093	31	AC011483	Homo sapiens chromosome 2	8.3e+02
19	130	41.1	166880	28	AC009756	Homo sapiens chromosome 4	3.17e+02
20	130	41.1	276261	21	HSAC000406	Homo sapiens chromosome 4	3.17e+02
21	127	40.2	10484	29	AC014515	Drosophila melanogaster	4.39e+02
22	127	40.2	17438	29	AC014821	Drosophila melanogaster	4.39e+02
23	127	40.2	119844	24	AC010044	Drosophila melanogaster	4.39e+02
24	127	40.2	121576	24	AC010043	Drosophila melanogaster	4.39e+02
25	127	40.2	143126	32	AC012603	Homo sapiens chromosome 4	3.9e+02
26	127	40.2	148476	24	AC023812	Homo sapiens chromosome 4	3.9e+02
27	127	40.2	162985	34	AC020751	Homo sapiens chromosome 4	3.9e+02
28	127	40.2	166290	30	AC010815	Homo sapiens chromosome 4	3.9e+02
29	127	40.2	223724	21	AP000834	Homo sapiens chromosome 4	3.9e+02
30	127	40.2	283187	31	AC010615	Homo sapiens chromosome 4	3.9e+02
31	125	39.6	71307	28	AC008296	Drosophila melanogaster	5.44e+02
32	125	39.6	83714	33	AC008297	Drosophila melanogaster	5.44e+02
33	125	39.6	97308	29	AC013925	Drosophila melanogaster	5.44e+02
34	125	39.6	131199	23	AC007807	Drosophila melanogaster	5.44e+02
35	125	39.6	132819	28	AC009335	Homo sapiens chromosome 5	4.4e+02
36	125	39.6	188432	34	AC018497	Homo sapiens chromosome 5	4.4e+02
37	122	38.6	86367	29	AC013310	Homo sapiens chromosome 5	4.4e+02
38	122	38.6	98750	34	AC020580	Homo sapiens chromosome 5	4.4e+02
39	122	38.6	126429	31	AC010447	Arabidopsis thaliana	7.45e+02
40	122	38.6	146461	31	AC010459	Homo sapiens chromosome 7	4.4e+02
41	122	38.6	145619	31	AC015474	Homo sapiens chromosome 7	4.4e+02
42	122	38.6	173737	33	AC022245	Homo sapiens chromosome 7	4.4e+02
43	122	38.6	175314	31	AC012653	Homo sapiens chromosome 7	4.4e+02
44	122	38.6	216858	32	AC008060	Homo sapiens chromosome 7	4.4e+02
45	122	38.6	228143	31	AC010422	Homo sapiens chromosome 7	4.4e+02

## ALIGNMENTS

RESULT	1	AC011523	40334 bp	DNA	HTG	07-OCT-1999
LOCUS		HOMO SAPIENS CHROMOSOME 19 CLONE LINT-F-214C7				*** SEQUENCING IN
DEFINITION		AC011523				PROGRESS *** 1 ordered pieces.
ACCESSION		AC011523.1	GI:6015229			
VERSION		HTG: HTGS_PHASE2				
KEYWORDS		human.				
SOURCE		Homo sapiens				
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE		1 (bases 1 to 40334)				
AUTHORS		DOE Joint Genome Institute.				
TITLE		Sequencing of Human Chromosome 19				
JOURNAL		Unpublished				
COMMENT		2 (bases 1 to 40334)				
		DOE Joint Genome Institute.				
		Direct Submission				
		Submitted (07-OCT-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA				
		www.jgi.doe.gov#mapping database ID 25479.				
		* NOTE: This is a "working draft" sequence. It currently				
		* consists of 1 contigs. Gaps between the contigs				
		* are represented as runs of N. The order of the pieces				
		* is believed to be correct as given, however the sizes				
		* of the gaps between them are based on estimates that have				
		* provided by the submitter.				
		* This sequence will be replaced				
		* by the finished sequence as soon as it is available and				
		* the accession number will be preserved.				
		1 40334: contig of 40334 bp in length.				
		Location/Qualifiers				
		1. 40334				
		/organism="Homo sapiens"				
		/db_xref="taxon:9606"				
		/chromosome="19"				

## FEATURES

source	
--------	--

BASE COUNT 9406 a 9604 c 10033 g 11291 t  
 ORIGIN /clone="LNL-F\_214C7"

Query Match 87.7%; Score 277; DB 28; Length 40334;  
 Best Local Similarity 59.7%; Pred. No. 7.32e-08;  
 Matches 40; Conservative 17; Mismatches 10; Indels 0; Gaps 0;

Db 26757 TCTTACCCCAAGAACTGAGTGTGACCTCATGTTATTTCCATGACGCTGTG 26816

QY 8 TTTTACCCCAAGAACTGAGTGTGACCTCATGTTATTTCCATGACGCTGTG 67

Db 26817 CGCAAGT 26823  
 QY 68 CMCARGT 74

RESULT 2 AC020019 40871 bp DNA HTG 03-JAN-2000  
 LOCUS Drosophila melanogaster, \*\*\* SEQUENCING IN PROGRESS \*\*\*  
 DEFINITION pieces.

ACCESSION AC020019.1 GI:6664878  
 VERSION HTG: HTGS\_PHASE2.  
 KEYWORDS fruit fly.  
 SOURCE Drosophila melanogaster

ORGANISM Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 1 (bases 1 to 40871)

REFERENCE Adams, M. and Venter, J.C.  
 AUTHORS Direct Submission  
 TITLE Submitted (30-DEC-1999) Celera Genomics, 45 West Gude Drive,  
 JOURNAL Rockville, MD, USA

COMMENT This sequence was identified as CDM:10211762 by the submitter.  
 For more information on this record e-mail to fly@celera.com.

\* NOTE: This is a 'working draft' sequence.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.  
 Location/Qualifiers

FEATURES  
 Source 1.40871  
 /organism="Drosophila melanogaster"  
 /db\_xref="taxon:7227"

BASE COUNT 12454 a 8202 c 8028 g 12187 t  
 ORIGIN

Query Match 45.9%; Score 145; DB 31; Length 40871;  
 Best Local Similarity 47.9%; Pred. No. 5.49e+01;  
 Matches 23; Conservative 12; Mismatches 13; Indels 0; Gaps 0;

Db 763 AAGAAATTTTAACTACACAAATGCTGTGTGATTTCTATGCA 810

QY 2 ARGARTTYTNACNCNAAARATTCARCTGTGATTCATGCA 49

RESULT 3 AC019591 2286 bp DNA HTG 03-JAN-2000  
 LOCUS Drosophila melanogaster, \*\*\* SEQUENCING IN PROGRESS \*\*\*  
 DEFINITION pieces.

ACCESSION AC019591.1 GI:6665306  
 VERSION HTG: HTGS\_PHASE2.  
 KEYWORDS fruit fly.  
 SOURCE Drosophila melanogaster

ORGANISM Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 1 (bases 1 to 2286)  
 REFERENCE Adams, M. and Venter, J.C.  
 AUTHORS Direct Submission  
 TITLE Submitted (30-DEC-1999) Celera Genomics, 45 West Gude Drive,  
 JOURNAL Rockville, MD, USA

## COMMENT

This sequence was identified as CDM:10210313 by the submitter.  
 For more information on this record e-mail to fly@celera.com.  
 \* NOTE: This is a 'working draft' sequence.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.  
 Location/Qualifiers

FEATURES  
 source 1.2286  
 /organism="Drosophila melanogaster"  
 /db\_xref="taxon:7227"

BASE COUNT 736 a 306 c 330 g 914 t  
 ORIGIN

Query Match 43.4%; Score 137; DB 31; Length 2286;  
 Best Local Similarity 42.4%; Pred. No. 1.43e+02;  
 Matches 25; Conservative 15; Mismatches 19; Indels 0; Gaps 0;

Db 1765 AAGATTTCTTATGTAAAGAAATTTGATACGAGATTCATTAAATTCGCAATAT 1823

QY 2 ARGARTTYTNACNCNAAARATTCARCTGTGATTCATGCA 60

RESULT 4 AC013501 60074 bp DNA HTG 13-NOV-1999  
 LOCUS Homo sapiens chromosome 3 clone 219\_A\_11 map 3, LOW-PASS SEQUENCE  
 DEFINITION SAMPLING.

ACCESSION AC013501.1 GI:6403780  
 VERSION HTG: HTGS\_PHASE0.  
 KEYWORDS human.  
 SOURCE Homo sapiens

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 1 (bases 1 to 60074)

REFERENCE Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
 AUTHORS Homo sapiens chromosome 3, clone 219\_A\_11  
 TITLE Unpublished  
 JOURNAL 2 (bases 1 to 60074)

REFERENCE Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
 AUTHORS Homo sapiens chromosome 3, clone 219\_A\_11  
 TITLE Unpublished  
 JOURNAL 2 (bases 1 to 60074)  
 COMMENT Submitted (13-NOV-1999) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 All repeats were identified using RepeatMasker. Smt, A.F.A. &  
 Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RN/RepeatMasker.html.  
 \* NOTE: This record contains 64 individual  
 \* sequencing reads that have not been assembled into  
 \* contigs. Runs of N are used to separate the reads  
 \* and the order in which they appear is completely  
 \* arbitrary. Low-pass sequence sampling is useful for  
 \* identifying clones that may be gene-rich and allows  
 \* overlap relationships among clones to be deduced.  
 \* However, it should not be assumed that this clone  
 \* will be sequenced to completion. In the event that  
 \* the record is updated, the accession number will  
 \* be preserved.

1 897: contig of 897 bp in length  
 \* gap of unknown length  
 \* 898 1891: contig of 994 bp in length  
 \* gap of unknown length









```

FEATURES
  source
    * 152212 164088: contig of 11877 bp in length
    * 164089 176591: contig of 12503 bp in length
    * 176592 189901: contig of 13310 bp in length
    * 189902 206879: contig of 16978 bp in length
    * 206880 228527: contig of 21648 bp in length.
      Location/Qualifiers
        1..228527
          /organism="Homo sapiens"
          /db_xref="taxon:9606"
          /clone_1fb="RP11-315E23"
          /clone_1lb="RPCT-11"

BASE COUNT      69299 a 46162 c 45464 g 67423 t   179 others
ORIGIN
Query Match      43.4%; Score 137; DB 22; Length 228527;
Best Local Similarity 43.5%; Pred. No. 1,43e+02;
Matches 30; Conservative 14; Mismatches 25; Indels 0; Gaps 0;

Db 60353 AGTAATTATTTAAATAAAGTGTCGGATGCATGATGTCCACATGYGTGAAGAATGAG 60412
Qy       2 ARGARTTYTNMACNCNMAARAARYTCARTGTGTGAYTTCATYGTNATHMSNAAYGAYG 61
Db 60413 TT TGAGCCC 60421
Qy       62 TTTGTGCNC 70
         | ||| |
         | ||| |

RESULT      8
LOCUS      AC008969 135499 bp DNA HTG 18-FEB-2000
DEFINITION Homo sapiens chromosome 19 clone CTD-2368P22, WORKING DRAFT
ACCESSION  AC008969
VERSION     AC008969.3 GI:6997052
KEYWORDS   HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE   1 (bases 1 to 135499)
            DOE Joint Genome Institute.
            Sequencing of Human Chromosome 19
            Unpublished
            2 (bases 1 to 135499)
            DOE Joint Genome Institute.
            Direct Submision
            Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
            On Feb 18, 2000 this sequence version replaced gi:600941.
COMMENT     -----Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            -----Summary Statistics
            Consensus quality: 98498 bases at least Q40
            Consensus quality: 108756 bases at least Q30
            Consensus quality: 112157 bases at least Q20
            Estimated insert size: 135499; sum-of-contigs estimation
            Estimated insert size: 98200; agarose-fp estimation
            -----
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 33 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *
            * 1 1019: contig of 1019 bp in length
            * gap of unknown length

```

1020	2453:	contig of 134 bp in length
2454	3528:	contig of 1075 bp in length
3529	4877:	contig of 1349 bp in length
4878	6468:	contig of 1591 bp in length
6469	7516:	contig of 1048 bp in length
7517	8906:	contig of 1390 bp in length
8907	10009:	contig of 1103 bp in length
10010	11066:	contig of 1055 bp in length
11065	12117:	contig of 1053 bp in length
12118	13197:	contig of 1080 bp in length
13198	14212:	contig of 1015 bp in length
14213	15289:	contig of 1077 bp in length
15290	17341:	contig of 2052 bp in length
17342	18414:	contig of 1073 bp in length
18415	20370:	contig of 1956 bp in length
20371	21408:	contig of 1038 bp in length
21409	22494:	contig of 1086 bp in length
22495	23508:	contig of 1014 bp in length
23509	25938:	contig of 2430 bp in length
25939	26949:	contig of 1011 bp in length
26950	27954:	contig of 1005 bp in length
27955	29537:	contig of 1583 bp in length
29538	30893:	contig of 1356 bp in length
30894	32142:	contig of 1249 bp in length
32143	33359:	contig of 1217 bp in length
33360	36842:	contig of 3483 bp in length
36843	43445:	contig of 6603 bp in length
43446	49982:	contig of 6537 bp in length
49983	56244:	contig of 6262 bp in length
56245	70785:	contig of 14541 bp in length
70786	97210:	contig of 26425 bp in length
97211	135499:	contig of 38285 bp in length

FEATURES	SOURCE	Location/Qualifiers
1.	135499	/organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="19" /clone="CTD-2368P22"
BASE COUNT	36143 a 33819 c 33220 g 33260 t	57 others
ORIGIN		
Query Match	42.7%	Score 135; DB 24; Length 135499;

Best Local Similarity 48.9%; Pred. No. 1,806+0;  
Matches 22; Conservative 11; Mismatches 12; Indels 0; Gaps 0.

Df 103871 CATGCAGCTTAGACACAGAAGTTTGTGAGTCACAGAAATCTT 103915  
|||:::||::||::||::||::||::||::||::||::||::||  
Cp 46 CRTGNARTCNACRCRCAITGNARITTTTTNGNCGINARRAATTCYT 2

RESULT	9	HTG	16-NOV-1999
LOCUS	AC014821	17438 bp	DNA
DEFINITION	Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, In Orderre		
ACCESSION	AC014821		
VERSION	AC014821.1	GI:6436514	
KEYWORDS	HTG; HTGS_PHASE2.		
SOURCE	fruit fly.		
ORGANISM	Drosophila melanogaster		

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (phases 1 to 17438)	Adams, M. and Venter, J.C.	Direct Submission	Submitted (16-NOV-1999)	Celera Genomics, 45 West Gude Drive, Rockville, MD, USA
				This sequence was identified as CDM:10210569 by the submitter.

COMMENTS  
This sequence was identified as CM:U0210569 by the submitter.  
For further information on this sequence e-mail to [fly@celera.com](mailto:fly@celera.com)  
\* NOTE: This is a 'working draft' sequence.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
Location/Qualifiers

```

BASE COUNT      5022 a      3904 c      3790 g      4722 t
ORIGIN           /organism="Drosophila melanogaster"
                  /db_xref="taxon:7227"

```

	Query Match	41.8%	Score 132;	DB 29;	length 17438;
	Best Local Similarity	45.1%;	Pred. No.	2.55e+02;	
	Matches	23;	Mismatches	15;	Indels 0; Gaps 0
		Conservative			
D6	16810	AATCTCAGTATGTCGACCTACACTGTATCACCAACACAAGTCCGCCCAA	16860		
	:     :	:	:	:	:
Gy	22	AATTNCRTGTGTNGAIIYTCATCATGTAHNTSSNAYAGAIVGTITGVCNCAR	72		

	RESULT	10				
LOCUS	AC012665	148162 bp	DNA	HMG	31-JAN-2000	
DEFINITION	Homo sapiens clone RP11-19B11, WORKING DRAFT SEQUENCE, 11 unordered reseq.					
ACCESSION	AC012665					
VERSION	AC012665.3	GI:6838945				
KEYWORDS	HTG; HTGS-PHASE1; HTGS-DRAFT.					
SOURCE	human.					
ORGANISM	Homo sapiens					

REFERENCE	AUTHORS	TITLE	JOURNAL	REFERENCE
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia		
		Eutheria; Primates; Carnivora; Homiidae; Homo.		
		1 (bases 1 to 148162)		
	Waterston, R.H.			
		The sequence of Homo sapiens clone		
		unpublished		
		2 (bases 1 to 148162)		

authors  
 title  
 JOURNAL  
 COMMENT  
 Direct Submission  
 Submitted (02-NOV-1999) Genome Sequencing Center, Washington  
 University School of Medicine, 4444 Forest Park Parkway, St. Louis  
 MO 63108, USA  
 On Jan 31, 2000 this sequence version replaced g1:6358866.

----- Genome Center -----  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC



[illegible]

Query Match	41.8%	Score 132;	DB 31;	Length 149x08;
Best Local Similarity	42.1%;	Pred.	No. 2.55e+02;	
Matches	24;	Conservative	14;	Mismatches 19; Indels 0; Gaps 0;
Db	87964	CTACACAGCAAGAAGAAACGCGCATGTGGCAGCGTTTCCTGTATATCCATGCATGTGTG	88020	
	:: ::     :: :::	: :  :  :	:: ::	
Oy	9	YNTNANCNCNNARARYTINCARTGTGTNGAYTTCATGATNTAITHSMNAAYGAIGINTG	65	
RESULT	12			

[illegible]

RESULT 13  
LOCUS AC008460 156492 bp DNA HTG 26-JAN-2000  
DEFINITION Homo sapiens chromosome 5 clone CTC-352D12, LOW-PASS SEQUENCE  
SAMPLING.  
ACCESSION AC008460  
VERSION AC008460.2 GI:6759081  
KEYWORDS HTG; HTGS\_PHASE0.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE  
AUTHORS DOE Joint Genome Institute.  
TITLE 1 (bases 1 to 156492)  
JOURNAL Sequencing of Human Chromosome 5  
REFERENCE  
AUTHORS DOE Joint Genome Institute.  
TITLE 2 (bases 1 to 156492)  
JOURNAL Unpublished  
DIRECT SUBMISSION  
Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
On Jan 26, 2000 this sequence version replaced gi:5686606.  
-----Genome Center  
Center: Joint Genome Institute  
Center Code: JGI  
Web site: <http://www.jgi.doe.gov>  
-----  
\* NOTE: This record contains 106 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.  
1  
\* 756: contig of 756 bp in length  
\* gap of unknown length  
\* 757 1086: contig of 330 bp in length  
\* gap of unknown length  
\* 1087 1739: contig of 653 bp in length  
\* gap of unknown length  
\* 1740 2862: contig of 1123 bp in length  
\* gap of unknown length  
\* 2863 3766: contig of 904 bp in length  
\* gap of unknown length  
\* 3767 4535: contig of 769 bp in length  
\* gap of unknown length  
\* 4536 5747: contig of 1212 bp in length  
\* gap of unknown length  
\* 5748 5900: contig of 153 bp in length  
\* gap of unknown length  
\* 5901 6346: contig of 446 bp in length  
\* gap of unknown length  
\* 6347 7132: contig of 786 bp in length  
\* gap of unknown length  
\* 7133 7576: contig of 444 bp in length  
\* gap of unknown length  
\* 7577 7799: contig of 223 bp in length  
\* gap of unknown length  
\* 7800 8393: contig of 594 bp in length  
\* gap of unknown length  
\* 8394 9096: contig of 703 bp in length  
\* gap of unknown length  
\* 9097 9994: contig of 898 bp in length  
\* gap of unknown length  
\* 9995 10596: contig of 602 bp in length  
\* gap of unknown length  
\* 10597 11598: contig of 1002 bp in length  
\* gap of unknown length

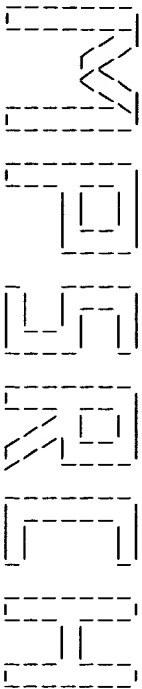
11599 12565: contig of 967 bp in length  
\* gap of unknown length  
\* 12566 13291: contig of 726 bp in length  
\* gap of unknown length  
\* 13292 14019: contig of 728 bp in length  
\* gap of unknown length  
\* 14020 14942: contig of 923 bp in length  
\* gap of unknown length  
\* 14943 15638: contig of 696 bp in length  
\* gap of unknown length  
\* 15639 16562: contig of 924 bp in length  
\* gap of unknown length  
\* 16563 17495: contig of 933 bp in length  
\* gap of unknown length  
\* 17496 18590: contig of 1095 bp in length  
\* gap of unknown length  
\* 18591 19370: contig of 780 bp in length  
\* gap of unknown length  
\* 19371 21061: contig of 1691 bp in length  
\* gap of unknown length  
\* 21062 21661: contig of 600 bp in length  
\* gap of unknown length  
\* 21662 22338: contig of 677 bp in length  
\* gap of unknown length  
\* 22339 23280: contig of 942 bp in length  
\* gap of unknown length  
\* 23281 24811: contig of 1531 bp in length  
\* gap of unknown length  
\* 24812 25884: contig of 1073 bp in length  
\* gap of unknown length  
\* 25885 26707: contig of 823 bp in length  
\* gap of unknown length  
\* 26708 27789: contig of 1062 bp in length  
\* gap of unknown length  
\* 27790 29117: contig of 1328 bp in length  
\* gap of unknown length  
\* 29118 30118: contig of 1001 bp in length  
\* gap of unknown length  
\* 30119 30860: contig of 742 bp in length  
\* gap of unknown length  
\* 30861 32125: contig of 1265 bp in length  
\* gap of unknown length  
\* 32126 33293: contig of 1168 bp in length  
\* gap of unknown length  
\* 33294 34626: contig of 1333 bp in length  
\* gap of unknown length  
\* 34627 35841: contig of 1215 bp in length  
\* gap of unknown length  
\* 35842 36803: contig of 962 bp in length  
\* gap of unknown length  
\* 36804 37840: contig of 1037 bp in length  
\* gap of unknown length  
\* 37841 38839: contig of 999 bp in length  
\* gap of unknown length  
\* 38840 39939: contig of 1100 bp in length  
\* gap of unknown length  
\* 39940 40973: contig of 1034 bp in length  
\* gap of unknown length  
\* 40974 41982: contig of 1009 bp in length  
\* gap of unknown length  
\* 41983 42857: contig of 875 bp in length  
\* gap of unknown length  
\* 42858 44198: contig of 1341 bp in length  
\* gap of unknown length  
\* 44199 45225: contig of 1027 bp in length  
\* gap of unknown length  
\* 45226 46646: contig of 1421 bp in length  
\* gap of unknown length  
\* 46647 48279: contig of 1633 bp in length  
\* gap of unknown length  
\* 48280 49444: contig of 1165 bp in length  
\* gap of unknown length  
\* 49445 51012: contig of 1568 bp in length





BASE COUNT	60653	a	37940	c	39345	g	64380	t	17	others
------------	-------	---	-------	---	-------	---	-------	---	----	--------

Search completed: Fri Jun 23 15:53:16 2000  
Job time : 299 secs.



(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPearch\_topn n.a. - n.a. Smith-Waterman search, using a protein query  
which has been backtranslated into n.a. using IUPAC symbols

Run on: Fri Jun 23 17:17:32 2000; Maspar time 177.79 Seconds  
1033.106 Million cell updates/sec

Tabular output not generated.

Title: >US-09-332-866-1  
Description: (1-25) from US09332866.pep  
Perfect Score: 316  
N.A. Sequence: 1 GARSGARTTYTNACNCNNA.....AYGAGTGTGCGCARGCTN 75  
Comp: CTCTCYAARRRANTGNGNTT.....TRCTRCANACGCGNGTTCAN

Scoring table: TABLE bkttranslated2  
Gap 40

Mmatch STD : Dbase 0; Query 0

Searched: 667327 seqs, 1224492533 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database:

genbank2-116  
1:gb-pat 2:gb-ph 3:gb-pl1 4:gb-pl2 5:gb-pl3 6:gb-pr1  
7:gb-pr2 8:gb-pr3 9:gb-pr4 10:gb-pr5 11:gb-pr6 12:gb-sts  
13:gb-sy 14:gb-un 15:gb-v1

Statistics: Mean 49.932; Variance 184.198; scale 0.271

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	312	98.7	990	6 HSN17040	Human prostate speci	2.23e-16
2	312	98.7	992	1 I67863	Sequence 13 from paten	2.23e-16
3	312	98.7	1415	6 HUMPA	Human prostate speci	2.23e-16
4	312	98.7	1446	6 HUMAPS	Human sapiens prostate	2.23e-16
5	312	98.7	1462	1 I67864	Sequence 14 from paten	2.23e-16
6	312	98.7	1466	7 HSPSAR	Human mRNA for prostat	2.23e-16
7	312	98.7	1654	6 HUMPA	Human prostate speci	2.23e-16
8	312	98.7	1729	7 HSPSA	Human mRNA for prostat	2.23e-16
9	312	98.7	1729	1 AR059540	Sequence 1 from patent	2.23e-16
10	282	89.2	1514	7 MPMROS	M.mulatta mRNA for pro	2.18e-13
11	277	87.7	281	7 HSPSA4	Human PSA gene for pro	6.78e-13
12	277	87.7	5873	7 HSPSAG	Human DNA for prostate	6.78e-13
13	277	87.7	6153	6 HUMPSANTIG	Human prostate-specifi	6.78e-13
14	277	87.7	7130	6 HUMPSAA	Human prostate speci	6.78e-13
15	277	87.7	7130	1 A37262	Sequence 2 from patent	6.78e-13
16	181	57.3	1341	9 S39529	glandular kallikrein-1	7.29e-04
17	181	57.3	2830	9 AF188746	Homo sapiens prostate	7.29e-04

18	181	57.3	2867	9 AF188747	Homo sapiens prostate	7.29e-04
19	180	57.0	569	9 S75755	PSA-prostate-specific	8.94e-04
20	161	50.9	15433	11 R4TRK5	R. norvegicus tonin (RS	3.96e-02
21	146	46.2	6139	7 HUMKAL2	Human glandular kallik	7.04e-01
22	145	45.9	858	6 MACKAL	Macaca fascicularis (C	8.50e-01
23	145	45.9	925	1 I09127	Sequence 2 from patent	8.50e-01
24	142	44.9	7400	15 HP01787	Human papillomavirus t	1.49e+00
25	141	44.6	863	11 PNRKALI	P. natalensis mRNA for	1.79e+00
26	137	43.4	115303	8 HSG12B18	Human DNA sequence fro	3.74e+00
27	135	42.7	658	6 HUMPA	Human prostate speci	5.39e+00
28	135	42.7	738	1 I08522	Sequence 2 from patent	5.39e+00
29	135	42.7	738	1 I02159	Sequence 7 from patent	5.39e+00
30	135	42.7	738	6 HUMKAL	Human kidney kallikrei	5.39e+00
31	135	42.7	800	7 BABRLK1SP	P. natalensis tissue	5.39e+00
32	135	42.7	827	11 PNGKALI	P. natalensis mRNA for	5.39e+00
33	135	42.7	870	11 MUSCPBM	Mouse major epididymal	5.39e+00
34	135	42.7	871	1 E01205	DNA sequence of human	5.39e+00
35	135	42.7	871	6 HUMKALX	Human kallikrein mRNA,	5.39e+00
36	132	41.8	73045	5 AC006218	Arabidopsis thaliana c	9.27e+00
37	132	41.8	157803	9 AC006256	Homo sapiens chromosom	9.27e+00
38	130	41.1	100906	4 ATF24G16	Arabidopsis thaliana D	1.33e+01
39	130	41.1	141190	3 AC003969	Homo sapiens chromosom	1.33e+01
40	127	40.2	2010	3 NT001961	Nicotiana tabacum call	2.26e+01
41	127	40.2	3341	3 ASTSSP12S	A. sativa 12S globulin	2.26e+01
42	127	40.2	7717	3 AP000380	Arabidopsis thaliana g	2.26e+01
43	127	40.2	8984	8 HUMPS	Homo sapiens prs gene,	2.26e+01
44	127	40.2	22980	7 HSL161A8	Human DNA from cosmid	2.26e+01
45	127	40.2	200348	9 AC005516	Homo sapiens Chromosom	2.26e+01

## ALIGNMENTS

RESULT	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	TITLE	REFERENCE	AUTHORS	JOURNAL	FEATURES
1	HSN17040	990 bp mRNA	U17040	U17040.1	GI:595945	human.	Homo sapiens	A Comparative Analysis of Prostate Specific Antigen Gene Sequence	Submitted (08-NOV-1994)	Jose M J G Moreno, Urology, Thomas Jefferson University Hospital, 1025 Walnut Street, Philadelphia, PA 19107, USA	Journal Unpublished	1.990 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="19" /map="19q13" /sex="male" /tissue-type="prostate" 42.827 /product="prostate specific antigen precursor" /protein_id="AAA56764.1" /db_xref="GI:595946" /translation="MMVPLVFLTSVTGAPLISRTVSGMECEKHSOPQVLYAS RGAVGVGVLPQVTPVAFHRLRKSVALIRSLRPHREDQGVFOYSHSPHPLVD MSLTKRFLRPDDSDSHDLMLRLSEPELTDVAVVMDLPDLPALGTCTASGWSGI EPSEFLTPKRLCCVDLHVTISNDVCAOVHOKVTKMLAGRWGKSTCSGDSGSPV CNGVLDGITSMSGSEPCALPERPSLYTKVHVHKKIKDIIVANP"

sig\_peptide



```

residue
/codon_start=1
/db_xref="GDB:G00-119-695"
/product="prostate-specific antigen"
/protein_id="AA38802.1"
/db_xref="GI:618464"
/translation="MWPFVFLTSLVWIGAAPLILSRIVGMECEKHSOPMOVLVAS
KRAVCGGVLPQWVLTAAHCIRNKSVILGRSLFPEPTGVFVSHSFPLPD
MSLKNRFLRPEDSSHDMLRLSEPAELDAYVMDLPTQEPALGTCTCAAGSGLI
EPEFLTPKLCQVDLHVISNDVCAQVHPQVTKFMLCAGRWGKSTCSGSGPLV
CNGVLGITSMSSEPCALPERSLTKVHYHRKMKITIVANP"
sig_peptide
25..75
/gene="APS"
/locus="G00-119-695"
/note="G00-119-695"
/mat_peptide
94..807
/gene="APS"
/note="G00-119-695"
/product="prostate-specific antigen"
3'UTR
811..1446
/gene="APS"
/note="G00-119-695"
/polyA_site
1446
/gene="APS"
/note="G00-119-695"
BASE COUNT 333 a 372 c 419 g 322 t
ORIGIN
Query Match 98.7% Score 312; DB 6; Length 1446;
Best Local Similarity 60.8%; Pred. No. 2,236-16;
Matches 45; Conservative 19; Mismatches 10; Indels 0; Gaps 0;
Db 511 GAGGAGTCTTGACCCCAAGAACTTCAGTGTGACCTCATGTATTCCAAATGAC 570
Oy 1 GARGARTTYTNACNCNAAARAARTNCARTGTGNGAYITNCAGTAAHKSNAAYGAY 60
Db 571 GGTGTGCGCAAGT 584
Oy 61 GTNTGYCNCARGT 74
LOCUS 5 167864 1462 bp DNA PAT 04-FEB-1998
DEFINITION Sequence 14 from patent US 5674682.
ACCESSION 167864
VERSION 167864.1 GI:2829986
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 1462)
Croce, C., Gomella, L., Mulholland, S. Grant, Moreno, J. G. and
Fischer, R.
Nucleic acid primers for detecting micrometastasis of prostate
cancer
Patent: US 5674682-A 14 OCT-1997;
JOURNAL Location/Qualifiers
FEATURES
1..1462
BASE COUNT 343 a 374 c 422 g 323 t
ORIGIN
Query Match 98.7% Score 312; DB 1; Length 1462;
Best Local Similarity 60.8%; Pred. No. 2,236-16;
Matches 45; Conservative 19; Mismatches 10; Indels 0; Gaps 0;
Db 511 GAGGAGTCTTGACCCCAAGAACTTCAGTGTGACCTCATGTATTCCAAATGAC 570
Oy 1 GARGARTTYTNACNCNAAARAARTNCARTGTGNGAYITNCAGTAAHKSNAAYGAY 60
Db 571 GGTGTGCGCAAGT 584
Oy 61 GTNTGYCNCARGT 74

```

```

RESULT 6
LOCUS HSPSAR 1466 bp mRNA PRI 21-MAR-1995
DEFINITION Human mRNA for prostate specific antigen.
ACCESSION X05332
VERSION X05332.1 GI:35740
KEYWORDS
antigen; kallikrein-like protein; prostate specific antigen; signal
peptide.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homidae; Homo.
REFERENCE
1 (bases 1 to 1415)
Lundvall, A. and Lilja, H.
Molecular cloning of human prostate specific antigen cDNA
JOURNAL FEBS Lett. 214 (2), 317-322 (1987)
MEDLINE 87190978
REFERENCE
2 (bases 1 to 1466)
Lundvall, A.
Direct Submission
Submitted (07-NOV-1989) to the EMBL/Genbank/DBJ databases
COMMENT Data kindly reviewed (01-DEC-1987) by Lundvall A.
FEATURES
source
1..1466
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue="prostate"
/clone_lib="(lambda)gt11"
/clone="lambda HPSA-1"
1..1466
/note="primary transcript"
44..829
/note="precursor polypeptide (AA -17 to 244)."
CDS
precursor_RNA
44..829
/note="precursor polypeptide (AA -17 to 244)."
/codon_start=1
/protein_id="CA28947.1"
/db_xref="GI:35741"
/db_xref="SWISS-PROT:P07288"
/translation="MWPFVFLTSLVWIGAAPLILSRIVGMECEKHSOPMOVLVAS
KRAVCGGVLPQWVLTAAHCIRNKSVILGRSLFPEPTGVFVSHSFPLPD
MSLKNRFLRPEDSSHDMLRLSEPAELDAYVMDLPTQEPALGTCTCAAGSGLI
EPEFLTPKLCQVDLHVISNDVCAQVHPQVTKFMLCAGRWGKSTCSGSGPLV
CNGVLGITSMSSEPCALPERSLTKVHYHRKMKITIVANP"
44..94
/product="signal peptide (AA -17 to -1)"
95..115
/note="propeptide (AA 1-7)"
116..826
/product="mature prostate specific antigen"
1445..1450
/note="put. polyA signal"
1466
/note="polyA site"
382 c 422 g 324 t
BASE COUNT 338 a 382 c 422 g 324 t
ORIGIN
Query Match 98.7% Score 312; DB 7; Length 1466;
Best Local Similarity 60.8%; Pred. No. 2,236-16;
Matches 45; Conservative 19; Mismatches 10; Indels 0; Gaps 0;
Db 530 GAGGAGTCTTGACCCCAAGAACTTCAGTGTGACCTCATGTATTCCAAATGAC 589
Oy 1 GARGARTTYTNACNCNAAARAARTNCARTGTGNGAYITNCAGTAAHKSNAAYGAY 60
Db 590 GGTGTGCGCAAGT 603
Oy 61 GTNTGYCNCARGT 74
RESULT 7
LOCUS HUMPA 1654 bp mRNA PRI 07-JAN-1995
DEFINITION Human prostate specific antigen (PA) gene, 3' end, clone PA 525.
ACCESSION M21896
VERSION M21896.1 GI:189525

```

KEYWORDS	glycoprotein; prostate antigen; prostate-specific antigen; semiogelin.
SOURCE	Human prostate tumor cell line PC 82, cDNA to mRNA, clone PA 525.
ORGANISM	Homo sapiens
REFERENCE AUTHORS	Eukariota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 1654) Riegman,P.H., Klaassen,P., van der Korput,J.A., Romijn,J.C. and Trapman,J. Molecular cloning and characterization of novel prostate antigen cDNA's Biochem. Biophys. Res. Commun. 155 (1), 181-188 (1989)
TITLE	
JOURNAL MEDLINE FEATURES	88326297
SOURCE	Location/Qualifiers 1..1654 /organism="Homo sapiens" /db_xref="taxon:9606" /map="19q13.3" <1..40 /gene="APS" /note="prostate specific antigen signal peptide" <1..707 /gene="APS" /note="prostate specific antigen precursor" /codon_start=2 /db_xref="GDB:600-119-695" /protein_id="AA59396.1" /db_xref="GI:189526" /translation="VFPLTISYTWGAAPLLISRIVSGECEKSHQPOVLYASGRNVCGGVLHPQWLVLTAAHCIRKNSVIILGRHSIFEDPGVOYSHSPFLYMSLIRKRLRPGDSSHDMLRLRSPEALDVAVMDLPTQEPALTGYASGSGSEEPFLEPKLOCCVDILNVVISNDYCAOVHQKTKTMLCAGRWGSKSTCSWILLTELMPRLPMVLHGSLVPWRGV" 1..706 /gene="APS" 62..703 /gene="APS" /note="prostate specific antigen"
BASE COUNT	353 a 468 C 454 G 379 T
ORIGIN	176 bp upstream of PvuII site.
Query Match	98.7%; Score 312; DB 6; Length 1654; Best local Similarity 60.8%; Pred. No.2.23e-16;
Matches	45; Conservative 19; Mismatches 10; Indels 0; Gaps 0;
Db	476 GAGGAGTTCTTGACCCCAAGAAGACTTCATGTGTGGACCTCATATTTCACATGAC 535   :     :  :     :  :     :  :     :  :     : Qy 1 GARGARTTYTNCCNCMAARAARYTNCARTGTGTGAYTTCAGTATMWSNAAYGAV 60
Db	536 GTGTGTGCCCACT 549      :       :
Qy	61 GTTGTGCNCARCT 74
RESULT	8
LOCUS	HSPSA 1729 bp mRNA PRI 27-MAR-1995
DEFINITION	Human mRNA for prostate specific antigen.
ACCESSION	X07730
VERSION	X07730.1 GI:35720
KEYWORDS	prostate specific antigen. human.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE AUTHORS	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 1729) Fittler,F. Direct Submission Submitted (19-MAY-1988) Fittler F., Institut fuer Physiologische Chemie der Universitaet, Schillerstr. 44, D-8000 Muenchen 2, FRG 2 (bases 1 to 1729) Schulz,P., Stucka,R., Feldmann,H., Combratio,G., Klobbeck,H.G. and Fittler,F. Sequence of a cDNA clone encompassing the complete mature human
TITLE	

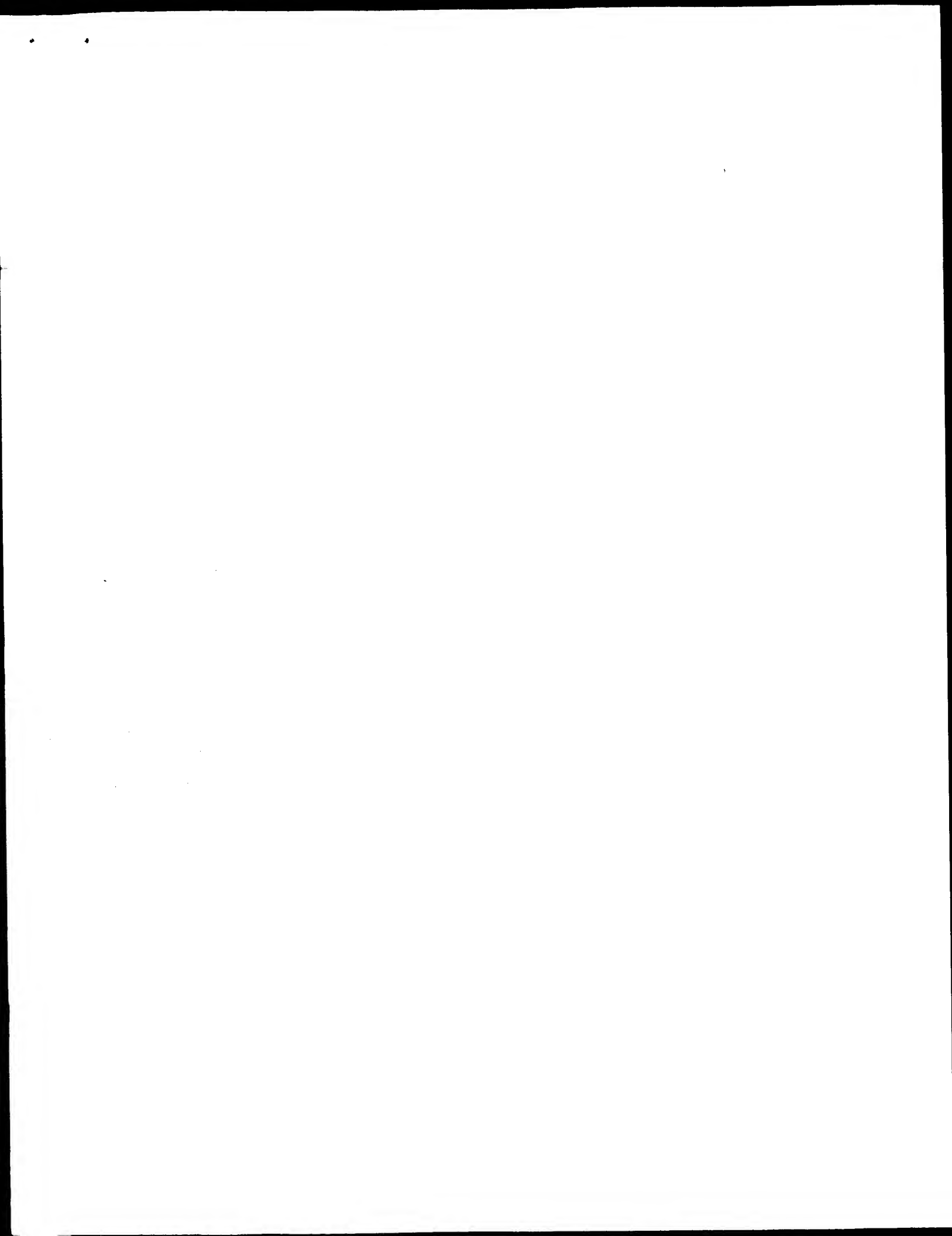
Journal	Medline	Comment	Features	Source	old_sequence	misc_feature	mat_peptide	misc_feature	BASE COUNT	Origin	Query Match	Best Local Similarity	Matches
JOURNAL	Prostate specific antigen (PSA) and an unspliced leader sequence												
Medline	Nucleic Acids Res. 16 (13), 6226 (1988)												
Comment	The sequence encompasses the complete mature human prostate specific antigen (PSA) and an unspliced leader												
Features	Data kindly reviewed (02-DEC-1988) by FITTLER F.												
Source	Location/Qualifiers												
	1. 1729												
	/organism="Homo sapiens"												
	/db_xref="taxon:9606"												
	/cell_line="LNCaP"												
	/clone_id="Lambda gtl1"												
	/note="t was ct in [2]"												
	4												
	/note="t was ct in [2]"												
	/citation=[2]												
	67. 69												
	/note="start codon"												
	378. 1088												
	/note="mature prostate specific antigen"												
	1706. 1711												
	/note="pot. polyA signal"												
	/note="pot. polyA signal"												
	449 g 383 t												
BASE COUNT	405 a 492 c 449 g 383 t												
Origin													
Query Match	98.7%; Score 312; DB 7; Length 1729;												
Best Local Similarity	60.8%; Pred. No. 2.23e-16;												
Matches	45; Conservative 19; Mismatches 10; Indels 0; Gaps 0;												
Db	792 GAGGAGTCTTGACCCCAAGAACTTCAGTGTGTGAGCCCTCCATGTATTTCAGTAC 851												
Qy	1 GARGARTTYTNACNCCNAARAAYTNCARTGTGTNAYTTCAYTNCAYGNATHWSNAAYGAY 60												
LOCUS	9 AR059540 1729 bp DNA PAT 29-SEP-1999												
DEFINITION	Sequence 1 from patent US 5840494.												
ACCESSION	AR059540												
VERSION	AR059540.1 GI:5985990												
KEYWORDS													
SOURCE	Unknown.												
ORGANISM	Unknown.												
REFERENCE	1 (bases 1 to 1729)												
AUTHORS	Katz, A.E., Buttyan, R., Rafio, A. and Olsson, C.A.												
TITLE	Method for molecular staging of prostate cancer												
JOURNAL	Patent: US 5840494-A 1 24-NOV-1998;												
FEATURES	Location/Qualifiers												
	1. 1729												
	/organism="unknown"												
BASE COUNT	405 a 492 c 449 g 383 t												









		/translation="MMVAVVFLTISVTWIGAAPILISRIYGNCEKHSOPQVIVAS RGRVAGGVIVHPQWVLTARAHCIKRSKVILLGRSLHPEDTQOVSHSPFHPYLD MSLKNRFLRPGDSSHDMLRLRSEPAELTDKAVNDLTPQEPALGTTCTVAGSGSI EPEEFELTPRKQLQCDVLHYISNDVCAQVHPQKVKFMICAGRWGKSTCGSGGPLY CNGVILQITISWSEPCALPERPSLYTKVHYRWIKDTIVANP" 401..446 /number=1 join(<401..446,1688..1847,3477..3763,3907..4043, 5413..>5568) 5413..446,1688..1713) sig_peptide join(401..446,1688..1713) 447..1687 /number=1 exon 1688..1847 /number=2 mat_peptide join(1714..1847,3477..3763,3907..4043,5413..5565) /product="prostate specific antigen (PSA)" 1848..3476 intron /number=2 misc_feature 2311..2663 /note="Alu repetitive sequence" 3477..3763 /number=3 exon 3764..3906 /number=3 intron 3907..4043 /number=4 exon 4044..5412 /number=4 exon 5413..>5568 /number=5 BASE COUNT 1166 a 1778 c 1503 g 1406 t ORIGIN	
Query Match Best Local Similarity 59.7%; Matches 40; Conservative 17; Mismatches 10; Indels 0; Gaps 0;		87.7%; Score 277; DB 7; Length 5873; Pred. No. 6,78e-13; Matches 40; Conservative 17; Mismatches 10; Indels 0; Gaps 0;	
Db 3907 TCTTGACCCCAAGAACTGAGTGTGAGACCTCCATGTTATTTCCATGACGTGTG 3966 QY 8 TTYTNACNCCNARARYNCARTGTGTGAYTNCAYTNATHMSNAAGAYGINTGYG 67			
Db 3967 CGCAAGT 3973 QY 68 CMCARGT 74			
RESULT 13 LOCUS HUMPSANTIG 6153 bp DNA PRI 08-JAN-1995 DEFINITION Human prostate-specific antigen (PA) gene, complete cds. ACCESSION M24543 VERSION M24543.1 GI:341200 KEYWORDS kallikrein; prostate-specific antigen. SOURCE Homo sapiens (tissue library: of G.grosveld) DNA. ORGANISM Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo. REFERENCE Riegman P.H., Vliestra,R.J., van der Korput,J.A., Romijn,J.C. and AUTHORS 1 (bases 1 to 6153) TITLE Characterization of the prostate-specific antigen gene: a novel JOURNAL human kallikrein-like gene MEDLINE Biochem. Biophys. Res. Commun. 159 (1), 95-102 (1989) FEATURES SOURCE 89165891 Location/Qualifiers 1..6153 /organism="Homo sapiens" /db_xref="taxon:9606" /tissue_lib="of G.grosveld" 71..84 /gene="PA" /note="transcriptional regulatory element" 188..193 /gene="PA"			
		TATA_signal 213..218 /gene="PA" exon 241..327 /gene="PA" /note="alternative exon 1" exon 247..327 /gene="PA" /note="alternative exon 1" gene join(282..327,1563..1725,3351..3637,3782..3918,5291.. CDS join(282..327,1563..1725,3351..3637,3782..3918,5291.. /gene="PA" /codon_start=1 /product="prostate-specific antigen" /protein_id="AAA60193.1" /db_xref="GI:511857" /translation="MMVAVVFLTISVTWIGAAPILISRIYGNCEKHSOPQVIVAS RGRVAGGVIVHPQWVLTARAHCIKRSKVILLGRSLHPEDTQOVSHSPFHPYLD MSLKNRFLRPGDSSHDMLRLRSEPAELTDKAVNDLTPQEPALGTTCTVAGSGMSI EPEEFELTPRKQLQCDVLHYISNDVCAQVHPQKVKFMICAGRWGKSTCGSGGPLY CNGVILQITISWSEPCALPERPSLYTKVHYRWIKDTIVANP" 328..1562 /gene="PA" intron /number=1 exon 1563..1725 /gene="PA" intron /number=2 exon 1727..3350 /gene="PA" intron /number=2 exon 3351..3637 /gene="PA" intron /number=3 exon 3638..3781 /gene="PA" intron /number=3 exon 3782..3918 /gene="PA" intron /number=4 exon 3919..5290 /gene="PA" intron /number=4 exon 5291..>6068 /gene="PA" polya_signal 6063..6068 /gene="PA" BASE COUNT 1277 a 1829 c 1592 g 1455 t ORIGIN	
Query Match Best Local Similarity 59.7%; Matches 40; Conservative 17; Mismatches 10; Indels 0; Gaps 0;		87.7%; Score 277; DB 6; Length 6153; Pred. No. 6,78e-13; Matches 40; Conservative 17; Mismatches 10; Indels 0; Gaps 0;	
Db 3782 TCTTGACCCCAAGAACTGAGTGTGAGACCTCCATGTTATTTCCATGACGTGTG 3841 QY 8 TTYTNACNCCNARARYNCARTGTGTGAYTNCAYTNATHMSNAAGAYGINTGYG 67			
Db 3842 CGCAAGT 3848 QY 68 CMCARGT 74			
RESULT 14 LOCUS HUMPSAA 7130 bp DNA PRI 08-JAN-1995 DEFINITION Human prostate specific antigen gene, complete cds. ACCESSION M27274 VERSION M27274.1 GI:190552 KEYWORDS Alu repeat; kallikrein; prostate-specific antigen. SOURCE Human peripheral leukocyte DNA, (library of R.Wydrö), clones 1ambda-PSA.124A.25A.25b1			











[illegible]

```

FT      Intron      1848..3476
FT      /*tag= 9
FT      /number= 2
FT      exon      3477..3763
FT      /*tag= h
FT      /number= 3
FT      Intron      3764..3906
FT      /*tag= 1
FT      exon      3907..4043
FT      /*tag= j
FT      /number= 4
FT      Intron      4044..5412
FT      /*tag= k
FT      /number= 4
FT      exon      5413..5568
FT      /*tag= 1
FT      /number= 5
WO9711172-A1.
PD      27-MAR-1997.
PF      20-SEP-1996: U15123.
PR      20-SEP-1995: US-004044.
(PORC-) WORCESTER FOUND BIOMEDICAL RES.
PI      Zamecnik PA;
DR      WPI: 97-202881/18.
P-PSDB: W13649.
PT      Treatment of benign prostatic hyperplasia or prostatic cancer
PT      using an anti-sense oligo:nucleotide targeted against prostate
PT      specific antigen or the probasin gene
PS      Claim 2: Page 23-30; 49pp: English.
CC      A novel method has been developed for treating a patient diagnosed as
CC      having benign prostatic hyperplasia or a prostatic cancer. The method
CC      comprises administering to the patient a therapeutic amount of a
CC      composition comprising an antisense oligonucleotide which selectively
CC      hybridises to prostate specific antigen (PSA) gene or mRNA or to a
CC      probasin gene or mRNA sequence of the patient, where the antisense
CC      oligonucleotide inhibits expression of the sequence. The present
CC      sequence encodes human PSA. The antisense oligonucleotide can be used
CC      to effectively inhibit the growth of (and for the killing of)
CC      hyperplastic cells or cancerous cells of prostatic origin. Expression
CC      of the PSA and probasin genes is specific to prostate cells so that the
CC      antisense oligonucleotide can be administered systemically, making it
CC      particularly useful in late stage prostatic cancer which has
CC      metastasised, and in which the cells have become resistant to oestrogen
CC      or anti-androgen therapy. The method can also be used in benign prostatic
CC      hyperplasia, or early stage prostatic cancer, as an alternative for the
CC      more radical procedures currently used, such as transurethral resection,
CC      radical prostatectomy, or physical or chemical castration.
CC      Sequence 5873 BP; 1186 A; 1778 C; 1503 G; 1406 T;
Query Match      87.7%; Score 277; DB 1; Length 5873;
Best Local Similarity 59.7%; Pred. No. 2,46e-12;
Matches 40; Conservative 17; Mismatches 10; Indels 0; Gaps 0;
Db 3907 TCTTGACCCCAAGAACTCAGTGTGTGACCTCATGTTATTTCCATGACGCTGTG 3966
QY 8 TTYTNACNCNNAARAARYTNCARTGTGTCNCAATYTNCAATGATGATGATGTCG 67
Db 3967 CGCAGCT 3973
QY 1 |||||
QY 68 CNCAGCT 74
RESULT 11
ID T90073 standard; cDNA: 90 BP.
AC T90073;
DT 14-APR-1998 (first entry)
DE Prostate specific antigen oligo-epitope encoding DNA.
KM Prostate specific antigen; PSA; epitope; vaccine; immune response;
KW Prostatic tumour cell; human leukocyte antigen; HLA; cytotoxic;
KW T lymphocyte; cancer; ds.
SS Synthetic.

```

```

FH Key Location/Qualifiers
FT misc_difference 58..60
FT tag a
FT /transl_except= (pos:58..60,aa:cys)
PN W09735021-A2.
PD 25-SEP-1997.
PF 19-MAR-1997: U04454.
PR 20-MAR-1996: US-618936.
PA (USSH ) US SEC DEPT HEALTH.
PI (USSH ) US DEPT HEALTH & HUMAN SERVICES.
DR Schlom J, Tsang K, Zaremba S;
DR P-PSDB: W27387.
PT Prostate specific antigen multiple epitope peptide - useful in
PT vaccine to produce immune response against prostate specific antigen
PT in prostatic tumor cell.
PS Claim 21: Page 64: 85pp. English.
CC The present sequence encodes a prostate specific antigen oligo-epitope
CC peptide (PSA-OP) of the present invention. The PSA-OP comprises more
CC than 1 adjoined PSA-OP which generates a prostate specific response in
CC a portion of the human population. The PSA-OP can be used to produce
CC PSA specific cytotoxic T lymphocytes, in vivo or in vitro, which can
CC kill or prevent the establishment or growth of PSA positive tumor
CC cells, i.e. to prevent or treat prostatic cancer. The products may be
CC used to elicit a PSA specific immune response. When a recombinant virus
CC is used to elicit the response it is coadministered with a carrier, a
CC biological response modifier e.g. interleukin 2, 6 or 12, interferon,
CC tumor necrosis factor, granulocyte/macrophage-colony stimulating factor
CC or cyclophosphamide and an adjuvant selected from Ribi Detox, QS21,
CC alum and incomplete Freund's adjuvant. PSA-OP comprises multiple PSA
CC epitopes, allowing an immune response to PSA in a population of humans
CC having more than 1 HLA class I molecule type to be generated.
SQ Sequence 90 BP; 23 A; 24 C; 19 G; 24 T;

Query Match 86.1%; Score 272; DB 1; Length 90;
Best Local Similarity 58.8%; Pred. No. 6,64e-12;
Matches 40; Conservative 17; Mismatches 11; Indels 0; Gaps 0;

DB 1 TTCTGACCCCAAGAACTTCAGTGTGAGCCTCCATGTTATTCACATGACGCT 60
QY 7 TTTTATNCNCNARARATNCARTGTGNGAYTNCATGATNATWSNATGAYGINTGT 66
DB 61 GCGCAGT 68
QY 67 GCMCAKGT 74

RESULT 12
ID V06602 standard; cDNA: 711 BP.
AC V06602:
DE Mature prostate-specific glandular kallikrein hk2 coding region.
DE Prostate-specific glandular kallikrein: hk2 protein; hK2 gene;
DE Prostate carcinoma; prostate cancer; benign prostatic hyperplasia;
DE antigen; diagnosis; human; ds.
OS Homo sapiens.
PN W09802748-A1.
PD 22-JAN-1998.
PF 15-JUL-1997: U12322.
PR 15-JUL-1996: US-680868.
PA (HYBR-) HYBRITTECH INC.
PA (MAYO-) MAYO FOUNDATION.
PI Grauer L, Klee GG, Mikolajczyk SD, Saedi M, Tindall DJ,
PI Young CYF;
DR WPI: 98-120378/11.
DR P-PSDB: W45395.
PT Diagnostic methods using antibodies which bind prostate antigens -
PT useful for, e.g. monitoring treatment or progression of prostate
PT cancer.
PS Example 1: Page 49-50; 100pp; English.
CC This cDNA comprises the coding region for human mature
CC prostate-specific glandular kallikrein hk2 (see W45395). cDNA
CC (see V06604) encoding the entire prepro-hk2 (pphk2, see W45397) was

```

```

CC synthesised from RNA of human benign prostate hyperplasia (BPH)
CC tissue by RT-PCR using hk2-specific primers (see V06605-06), and
CC inserted into vector pVIL1393. Plasmid pphk2/pVIL1393 in E. coli
CC HB101 is deposited as ATCC 69614. The cDNA for the mature hk2
CC peptide can be utilised in the production of recombinant hk2 in
CC host cells. The invention relates to diagnostic methods employing
CC antibodies which bind to prostate antigen hk2 or pro-hk2 (see also
CC W45396) and which do not react significantly with prostate specific
CC antigen hk3 (see W45398). The methods are useful for the diagnosis
CC of prostate cancer and for monitoring the treatment and/or
CC progression, or the early detection, of prostate cancer in a male
CC who has BPH or a high grade prostatic neoplasia (HPGN) or a male
CC whose family members have or had BPH, HPGN or a prostate cancer.
SQ Sequence 711 BP; 153 A; 195 C; 210 G; 153 T;

Query Match 57.3%; Score 181; DB 1; Length 711;
Best Local Similarity 47.1%; Pred. No. 2.57e-04;
Matches 32; Conservative 16; Mismatches 20; Indels 0; Gaps 0;

DB 415 GAGGAGTTCTGCGCCCGAGAGTTCAGTGTGAGCCTCCATGCTCTCCATGAC 474
QY 1 GAGGARTTTTNCNCNARARATNCARTGTGNGAYTNCATGATNATWSNATGAY 60
DB 475 ATGTGTGC 482
QY 61 GTMTGYGC 68

RESULT 13
ID V70339 standard; cDNA: 711 BP.
AC V70339:
DE 09-FEB-1999 (first entry)
DE Prostate-specific glandular kallikrein protein hk2 encoding cDNA.
DE Prostate cancer; detection; hk2; hK1; hK3; pphk2; tissue kallikrein;
DE pphk2; prostate-specific glandular kallikrein protein; PSA; human;
DE prostate-specific antigen; ss.
OS Homo sapiens.
PN Slawin KM, Tindall DJ, Young CYF;
DR WPI: 98-594592/50.
DR P-PSDB: W83202.
PT Detection of human kallikrein 2 RNA - by reverse transcription and
PT amplification by PCR, for detecting, monitoring and staging of
PT prostate cancer.
PS Example 1: Fig 2; 90pp; English.
CC The present invention describes a diagnostic method for detecting human
CC kallikrein 2 (hk2) DNA. The method comprises: (a) contacting DNA obtained
CC by reverse transcription (RT) of RNA from a human physiological sample
CC which comprises cells suspected of containing hk2 RNA with at least 2
CC oligonucleotides to amplify the DNA by PCR to yield amplified hk2 DNA,
CC where the conditions amplify the DNA obtained by RT of RNA from at least
CC one cell containing hk2 in a sample which comprises at least 107 to 109
CC cells; and (b) detecting the presence of the amplified hk2 DNA. The
CC method can be used for detecting, monitoring the progression of and
CC pathologically staging prostate cancer. The present sequence encodes
CC prostate-specific glandular kallikrein protein (hk2).
SQ Sequence 711 BP; 153 A; 195 C; 210 G; 153 T;

Query Match 57.3%; Score 181; DB 1; Length 711;
Best Local Similarity 47.1%; Pred. No. 2.57e-04;
Matches 32; Conservative 16; Mismatches 20; Indels 0; Gaps 0;

DB 415 GAGGAGTTCTGCGCCCGAGAGTTCAGTGTGAGCCTCCATGCTCTCCATGAC 474

```



QY 1 GARGARTTYTNACNCNARARATNCARTGCTNGAYTNCAYGTNATHWSNAYGAY 60  
 Db 475 ATGTGTGC 482  
 QY 61 GTNTGYGC 68

## RESULT 14

ID T05148 standard; cDNA; 754 BP.  
 AC T05148;

DT 04-MAR-1996 (first entry)  
 DE kallikrein hk2 cDNA.

KW kallikrein hk2; serine protease; prostate cancer; diagnosis; ds.  
 OS Homo sapiens.

PH key Location/Qualifiers  
 FT CDS 7..723

FT MO9530758-A1. /\*tag= a

PD 16-NOV-1995.

PF 09-MAY-1995; U06157.

PR 10-MAY-1994; US-241174.

PR 02-MAY-1995; US-427767.

PA (HYBR-) HYBRITTECH INC.

PI (MAYO-) MAYO FOUNDATION.

PI Saeed MS, Tindall DJ, Young CYF;

DR P-PSDB; R84670.

PT New isolated prostate-specific kallikrein polypeptide(s) - used to develop prods. for use in assays for such polypeptide(s), partic.

PS Disclosure: Page 35-36; 61pp; English.

CC cDNA (T05148) coding for human hk2 kallikrein (R84670) is obt. by PCR amplification of prepro-hk2 cDNA (T05147) using hk2-specific

CC primers (T05150, T05152). The cDNA is used for large-scale prodn. in prokaryotic (E. coli) or eukaryotic (insect, mammalian) cells of

CC recombinant hk2, a prostate-specific glandular kallikrein.

CC Sequence 754 BP; 164 A; 211 C; 215 G; 164 T;

Query Match 57.3%; Score 181; DB 1; Length 754;  
 Best Local Similarity 47.1%; Pred. No. 2.57e-04;

Matches 32; Conservative 16; Mismatches 20; Indels 0; Gaps 0;

Db 424 GAGGAGTTCCTGCGCCCGAGAGTCTTCAGTGTGAGCCCTCCCTGCTCCATGAC 483  
 QY 1 GARGARTTYTNACNCNARARATNCARTGCTNGAYTNCAYGTNATHWSNAYGAY 60

Db 484 ATGTGTGC 491  
 QY 61 GTNTGYGC 68

Db 484 ATGTGTGC 491  
 QY 61 GTNTGYGC 68

Db 484 ATGTGTGC 491  
 QY 61 GTNTGYGC 68

Db 484 ATGTGTGC 491  
 QY 61 GTNTGYGC 68

Db 484 ATGTGTGC 491  
 QY 61 GTNTGYGC 68

Db 484 ATGTGTGC 491  
 QY 61 GTNTGYGC 68

Db 484 ATGTGTGC 491  
 QY 61 GTNTGYGC 68

Db 484 ATGTGTGC 491  
 QY 61 GTNTGYGC 68

Db 484 ATGTGTGC 491  
 QY 61 GTNTGYGC 68

Db 484 ATGTGTGC 491  
 QY 61 GTNTGYGC 68

Db 484 ATGTGTGC 491  
 QY 61 GTNTGYGC 68

Db 484 ATGTGTGC 491  
 QY 61 GTNTGYGC 68

Db 484 ATGTGTGC 491  
 QY 61 GTNTGYGC 68

Db 484 ATGTGTGC 491  
 QY 61 GTNTGYGC 68

Db 484 ATGTGTGC 491  
 QY 61 GTNTGYGC 68

Db 484 ATGTGTGC 491  
 QY 61 GTNTGYGC 68

Db 484 ATGTGTGC 491  
 QY 61 GTNTGYGC 68

Db 484 ATGTGTGC 491  
 QY 61 GTNTGYGC 68

Db 484 ATGTGTGC 491  
 QY 61 GTNTGYGC 68

Db 484 ATGTGTGC 491  
 QY 61 GTNTGYGC 68

Db 484 ATGTGTGC 491  
 QY 61 GTNTGYGC 68

Db 484 ATGTGTGC 491  
 QY 61 GTNTGYGC 68

Db 484 ATGTGTGC 491  
 QY 61 GTNTGYGC 68

PT amplification by PCR, for detecting, monitoring and staging of prostate cancer

PS Example 1; Fig 2; 90pp; English.

CC The present invention describes a diagnostic method for detecting human kallikrein 2 (hk2) DNA. The method comprises: (a) contacting DNA obtained

CC by reverse transcription (RT) of RNA from a human physiological sample which comprises cells suspected of containing hk2 RNA with at least 2

CC oligonucleotides to amplify the DNA by PCR to yield amplified hk2 DNA, where the conditions amplify the DNA obtained by RT of RNA from at least

CC one cell containing hk2 in a sample which comprises at least 107 to 109 cells; and (b) detecting the presence of the amplified hk2 DNA. The

CC method can be used for detecting, monitoring the progression of and pathologically staging prostate cancer. The present sequence encodes

CC prostate-specific glandular kallikrein protein (phk2).

CC Sequence 766 BP; 162 A; 218 C; 220 G; 166 T;

Query Match 57.3%; Score 181; DB 1; Length 766;  
 Best Local Similarity 47.1%; Pred. No. 2.57e-04;

Matches 32; Conservative 16; Mismatches 20; Indels 0; Gaps 0;

Db 436 GAGGAGTTCCTGCGCCCGAGAGTCTTCAGTGTGAGCCCTCCCTGCTCCATGAC 495  
 QY 1 GARGARTTYTNACNCNARARATNCARTGCTNGAYTNCAYGTNATHWSNAYGAY 60

Db 496 ATGTGTGC 503  
 QY 61 GTNTGYGC 68

Db 496 ATGTGTGC 503  
 QY 61 GTNTGYGC 68

Db 496 ATGTGTGC 503  
 QY 61 GTNTGYGC 68

Db 496 ATGTGTGC 503  
 QY 61 GTNTGYGC 68

Db 496 ATGTGTGC 503  
 QY 61 GTNTGYGC 68

Db 496 ATGTGTGC 503  
 QY 61 GTNTGYGC 68

Db 496 ATGTGTGC 503  
 QY 61 GTNTGYGC 68

Db 496 ATGTGTGC 503  
 QY 61 GTNTGYGC 68

Db 496 ATGTGTGC 503  
 QY 61 GTNTGYGC 68

Db 496 ATGTGTGC 503  
 QY 61 GTNTGYGC 68

Db 496 ATGTGTGC 503  
 QY 61 GTNTGYGC 68

Db 496 ATGTGTGC 503  
 QY 61 GTNTGYGC 68

Db 496 ATGTGTGC 503  
 QY 61 GTNTGYGC 68

Db 496 ATGTGTGC 503  
 QY 61 GTNTGYGC 68

Db 496 ATGTGTGC 503  
 QY 61 GTNTGYGC 68

Db 496 ATGTGTGC 503  
 QY 61 GTNTGYGC 68

Db 496 ATGTGTGC 503  
 QY 61 GTNTGYGC 68

Db 496 ATGTGTGC 503  
 QY 61 GTNTGYGC 68

Db 496 ATGTGTGC 503  
 QY 61 GTNTGYGC 68

Db 496 ATGTGTGC 503  
 QY 61 GTNTGYGC 68

Db 496 ATGTGTGC 503  
 QY 61 GTNTGYGC 68

Db 496 ATGTGTGC 503  
 QY 61 GTNTGYGC 68

Db 496 ATGTGTGC 503  
 QY 61 GTNTGYGC 68

Db 496 ATGTGTGC 503  
 QY 61 GTNTGYGC 68

Db 496 ATGTGTGC 503  
 QY 61 GTNTGYGC 68

Db 496 ATGTGTGC 503  
 QY 61 GTNTGYGC 68

Db 496 ATGTGTGC 503  
 QY 61 GTNTGYGC 68

Db 496 ATGTGTGC 503  
 QY 61 GTNTGYGC 68

Db 496 ATGTGTGC 503  
 QY 61 GTNTGYGC 68

Db 496 ATGTGTGC 503  
 QY 61 GTNTGYGC 68

Db 496 ATGTGTGC 503  
 QY 61 GTNTGYGC 68

Db 496 ATGTGTGC 503  
 QY 61 GTNTGYGC 68

Db 496 ATGTGTGC 503  
 QY 61 GTNTGYGC 68

Db 496 ATGTGTGC 503  
 QY 61 GTNTGYGC 68

Db 496 ATGTGTGC 503  
 QY 61 GTNTGYGC 68

Db 496 ATGTGTGC 503  
 QY 61 GTNTGYGC 68

Db 496 ATGTGTGC 503  
 QY 61 GTNTGYGC 68

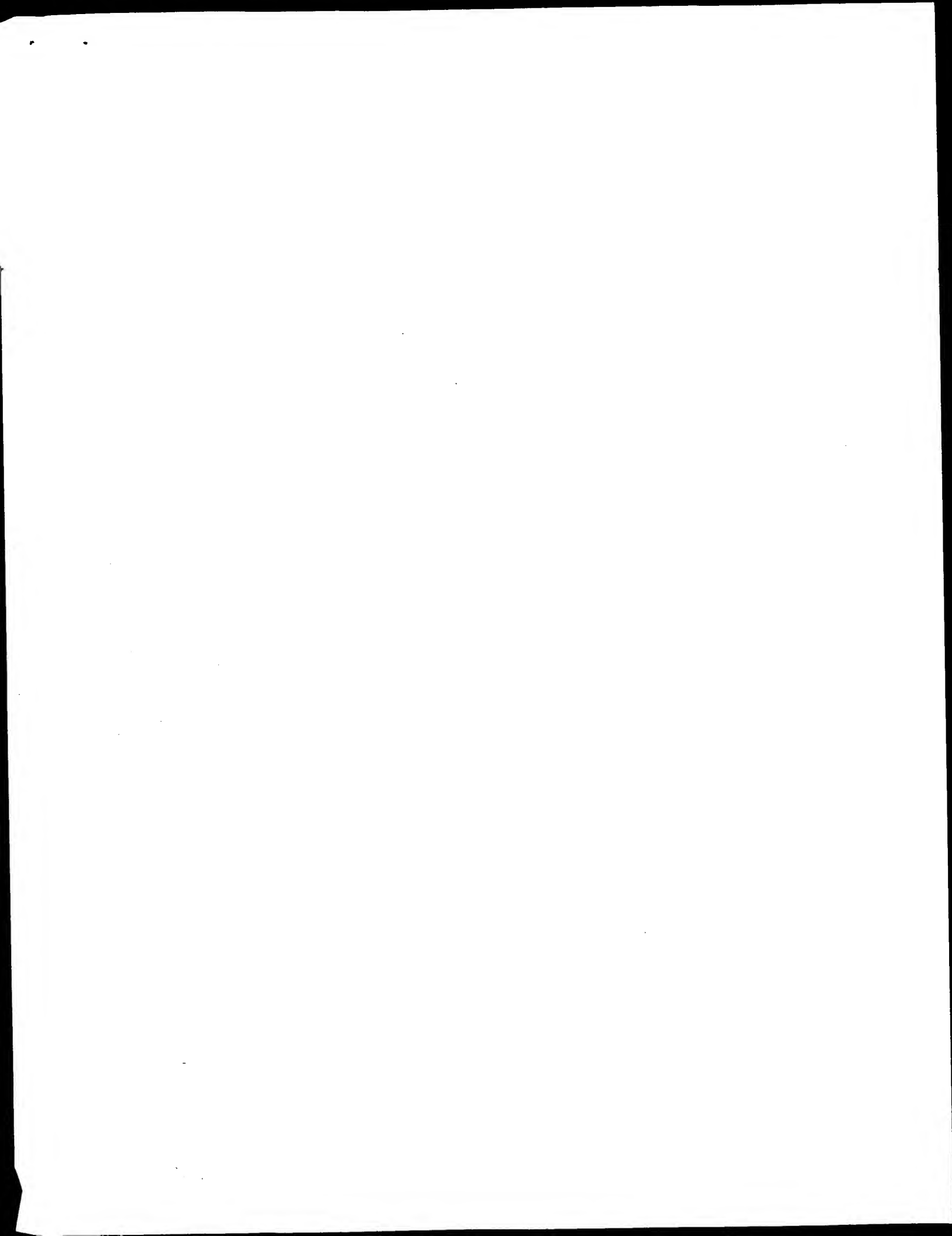
Db 496 ATGTGTGC 503  
 QY 61 GTNTGYGC 68

PT Detection of human kallikrein 2 RNA - by reverse transcription and

PT Detection of human kallikrein 2 RNA - by reverse transcription and

PT Detection of human kallikrein 2 RNA - by reverse transcription and

PT Detection of human kallikrein 2 RNA - by reverse transcription and



\*\*\*\*\*  
MUSEUM  
(TM)  
\*\*\*\*\*

Release 3.1a John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPorch\_tpm n.a. - n.a. Smith-Waterman search, using a protein query  
which has been backtranslated into n.a. using IUPAC symbols

Run on: Fri Jun 23 17:11:42 2000; Maspar time 15.09 Seconds  
645.987 Million cell updates/sec  
Tabular output not generated.

Title: >US-09-332-866-1  
Description: (1-25) from US09332866.pep  
Perfect Score: 316  
N.A. Sequence: 1 GARGARTTYTACNCNNA.....AYGATNTGYGCNCARGTN 75  
Comp: CTCTCAARANTGNGNT.....TTCRCNACNCGNGTTCAN

Scoring table: TABLE bkttranslated2  
Gap 40

Mmatch STD : Dbase 0: Query 0

Searched: 230473 seqs, 64993002 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: n-15used  
1:5A\_COMB 2:5B\_COMB 3:5C\_COMB 4:5D\_COMB 5:6\_COMB  
6:PCT\_COMB 7:backfile1

Statistics: Mean 37.556; Variance 145.681; scale 0.258

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	312	98.7	992	4	US-08-764-	Sequence 13, Applicat	1.18e-18
2	312	98.7	992	1	US-08-358-	Sequence 13, Applicat	1.18e-18
3	312	98.7	1462	1	US-08-358-	Sequence 14, Applicat	1.18e-18
4	312	98.7	1462	4	US-08-764-	Sequence 14, Applicat	1.18e-18
5	312	98.7	1729	4	US-08-764-	Sequence 1, Applicat	1.18e-18
6	312	98.7	1729	3	US-08-844-	Sequence 1, Applicat	1.18e-18
7	181	57.3	760	5	US-08-768-	Sequence 7, Applicat	5.30e-06
8	181	57.3	760	6	PCT-US95-0	Sequence 7, Applicat	5.30e-06
9	181	57.3	766	5	PCT-US95-0	Sequence 9, Applicat	5.30e-06
10	181	57.3	766	5	US-08-768-	Sequence 9, Applicat	5.30e-06
11	181	57.3	832	5	US-08-768-	Sequence 5, Applicat	5.30e-06
12	181	57.3	832	6	PCT-US95-0	Sequence 5, Applicat	5.30e-06
13	181	57.3	832	5	US-08-768-	Sequence 20, Applicat	5.30e-06
14	160	50.6	60	4	US-08-832-	Sequence 8, Applicat	4.20e-04
15	120	38.0	2815	4	US-08-940-	Sequence 1, Applicat	1.13e+00
16	120	38.0	2815	1	US-08-230-	Sequence 1, Applicat	1.13e+00
17	120	38.0	2815	2	US-08-619-	Sequence 1, Applicat	1.13e+00
18	110	34.8	396	1	US-08-470-	Sequence 43, Applicat	7.23e+00

19	110	34.8	396	1	US-08-470-	Sequence 121, Applicat	7.23e+00
20	110	34.8	1534	1	US-07-973-	Sequence 2, Applicat	7.23e+00
21	107	33.9	8067	2	US-08-078-	Sequence 1, Applicat	1.25e+01
22	107	33.9	8067	2	US-08-486-	Sequence 8, Applicat	1.25e+01
23	107	33.9	8067	1	US-08-117-	Sequence 8, Applicat	1.25e+01
24	105	33.2	5502	5	US-08-836-	Sequence 1, Applicat	1.78e+01
25	102	32.3	702	1	US-08-518-	Sequence 1, Applicat	3.04e+01
26	102	32.3	1209	6	PCT-US95-0	Sequence 5, Applicat	3.04e+01
27	102	32.3	1213	6	PCT-US95-0	Sequence 10, Applicat	3.04e+01
28	102	32.3	2540	4	US-08-231-	Sequence 15, Applicat	3.04e+01
29	102	32.3	2540	4	US-08-486-	Sequence 15, Applicat	3.04e+01
30	102	32.3	3007	4	US-08-480-	Sequence 15, Applicat	3.04e+01
31	102	32.3	3070	5	US-08-486-	Sequence 15, Applicat	3.04e+01
32	102	32.3	3070	5	US-08-480-	Sequence 33, Applicat	3.04e+01
33	102	32.3	3070	5	US-08-486-	Sequence 39, Applicat	3.04e+01
34	102	32.3	3211	5	US-08-480-	Sequence 39, Applicat	3.04e+01
35	102	32.3	3211	4	US-08-486-	Sequence 21, Applicat	3.04e+01
36	102	32.3	3274	4	US-08-486-	Sequence 21, Applicat	3.04e+01
37	102	32.3	3731	4	US-08-486-	Sequence 37, Applicat	3.04e+01
38	102	32.3	3731	5	US-08-480-	Sequence 31, Applicat	3.04e+01
39	102	32.3	3935	3	US-08-231-	Sequence 19, Applicat	3.04e+01
40	102	32.3	4094	5	US-08-480-	Sequence 29, Applicat	3.04e+01
41	102	32.3	4157	5	US-08-486-	Sequence 25, Applicat	3.04e+01
42	102	32.3	4157	4	US-08-486-	Sequence 25, Applicat	3.04e+01
43	102	32.3	4234	2	US-08-805-	Sequence 1, Applicat	3.04e+01
44	102	32.3	4234	1	US-08-486-	Sequence 1, Applicat	3.04e+01
45	102	32.3	4361	3	US-08-231-	Sequence 23, Applicat	3.04e+01

## ALIGNMENTS

RESULT 1  
ID US-08-764-527A-13 STANDARD; DNA; UNC; 992 BP.  
AC xxxxxx  
DE Sequence 13, Application US/08764527A  
CC Sequence 13, Application US/08764527A  
CC Patent No. 5939258  
CC GENERAL INFORMATION:  
CC APPLICANT: Cioce, Carlo  
CC APPLICANT: Gomella, Leonard  
CC APPLICANT: Mulholland, S. Grant  
CC APPLICANT: Moreno, Jose  
CC APPLICANT: Fischer, Rainer  
CC TITLE OF INVENTION: Methods of Detecting Micrometastasis of  
CC NUMBER OF SEQUENCES: 14  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESS: Woodcock Washburn Kurtz Mackiewicz &  
CC STREET: One Liberty Place 46th. Floor  
CC CITY: Philadelphia  
CC STATE: PA  
CC ZIP: 19103  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/764,527A  
CC FILING DATE:  
CC CLASSIFICATION:  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/358,782  
CC FILING DATE: 15-DEC-1994  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Beardsell, Lori Y.  
CC REGISTRATION NUMBER: 34,293  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 215-568-3100  
CC TELEFAX: 215-568-3439  
CC INFORMATION FOR SEQ ID NO: 13:



```

ID US-08-764-527A-14 STANDARD; DNA: UNC; 1462 BP.
AC xxxxxx
DT
DE Sequence 14, Application US/08764527A
CC Sequence 14, Application US/08764527A
CC Patent No. 5939258
CC GENERAL INFORMATION:
CC APPLICANT: Croce, Carlo
CC APPLICANT: Gomella, Leonard
CC APPLICANT: Mulholland, S. Grant
CC APPLICANT: Moreno, Jose
CC APPLICANT: Fischer, Rainer
CC TITLE OF INVENTION: Methods of Detecting Micrometastasis of
CC TITLE OF INVENTION: Prostate
CC NUMBER OF SEQUENCES: 14
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
CC STREET: One Liberty Place 46th. Floor
CC CITY: Philadelphia
CC STATE: PA
CC ZIP: 19103
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/764,527A
CC FILING DATE:
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US/08/358,782
CC FILING DATE: 15-DEC-1994
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Beardell, Lori Y.
CC REGISTRATION NUMBER: 34,293
CC REFERENCE/DOCKET NUMBER: JUU-1327
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 215-568-3100
CC TELEFAX: 215-568-3439
CC INFORMATION FOR SEQ ID NO: 14:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1462 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
SO SEQUENCE 1462 BP; 343 A; 374 C; 422 G; 323 T; 0 OTHER.

Query Match 98.7%; Score 312; DB 4; Length 1462;
Best Local Similarity 90.8%; Pred. No.1,18e-18;
Matches 45; Conservative 19; Mismatches 10; Indels 0; Gaps 0;

Db 511 GAGAGAGTCTTGACCCCAAGAAACTGCAGTGTGTGACGCTCATGTTATTTCCATGAC 570
Oy 1 GARGARTTYTNNACCCNAARARATYTCARCTGYGNGAYTNCAYGTNATHMSAAAGAY 60
Db 571 GTGTGTGCGCAAGT 584
Oy 61 GTNTGCGCCARGT 74

RESULT 5
ID US-08-718-547-1 STANDARD; DNA: UNC; 1729 BP.
AC xxxxxx
DT
DE Sequence 1, Application US/08718547
CC Sequence 1, Application US/08718547
CC Patent No. 5976794
CC GENERAL INFORMATION:
CC APPLICANT: Katz, Aaron E.; Buttyan, Ralph; Raifo, Anthony;

```

```

CC APPLICANT: Olsson, Carl A.
CC TITLE OF INVENTION: A Method For Molecular Staging Of
CC TITLE OF INVENTION: Prostate Cancer
CC NUMBER OF SEQUENCES: 4
CC CORRESPONDENCE ADDRESSES:
CC ADDRESSEE: Cooper & Dunham LLP
CC STREET: 1185 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: U.S.A.
CC ZIP: 10036
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/718,547
CC FILING DATE:
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: White, John P.
CC REGISTRATION NUMBER: 28,678
CC REFERENCE/DOCKET NUMBER: 43677-A-PCT/JPW/MJG
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 278-0400
CC TELEFAX: (212) 391-0525
CC TELEX:
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1729 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC ORIGINAL SOURCE:
CC ORGANISM: Homo sapiens
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 378..1088
CC SEQUENCE 1729 BP; 405 A; 492 C; 449 G; 383 T; 0 OTHER.
SQ
Query Match 98.7%; Score 312; DB 4; Length 1729;
Best Local Similarity 60.8%; Pred. No. 1,18e-18;
Matches 45; Conservative 19; Mismatches 10; Indels 0; Gaps 0;
Db 792 GAGGAGTCTCTGACCACCAAGAAACTTCAGTGTGACGCTCCATGTTATTTCCAAATGAC 851
OY 1 GARGATTYYTMAACCCCAARAAARATYTCARTGCTGYNAGTYTNCAYGTNATWSNAAAGAY 60
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 852 GTGTGTGCCCAAGT 865
|||:|||||:||||
OY 61 GTNTGYGCNCARGT 74
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
RESULT 6
ID US-08-844-024-1 STANDARD; DNA; UNC; 1729 BP.
AC xxxxxx
DT
DE Sequence 1, Application US/08844024
CC Sequence 1, Application US/08844024
CC Patent No. 5840494
CC GENERAL INFORMATION:
CC APPLICANT: Katz, Aaron E., et al.
CC TITLE OF INVENTION: A Method For Molecular Staging Of
CC TITLE OF INVENTION: Prostate Cancer
CC NUMBER OF SEQUENCES: 4
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Cooper & Dunham
CC STREET: 30 Rockefeller Plaza
CC CITY: New York
CC STATE: New York
CC COUNTRY: U.S.A.

```









```
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 60-014929
CC FILING DATE: 05-APR-1996
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Jay, Jeremy M.
CC REGISTRATION NUMBER: 33587
CC REFERENCE/DOCKET NUMBER: 72466
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 202-737-6770
CC TELEFAX: 202-737-6776
CC INFORMATION FOR SEQ ID NO: 8:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 60 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: other nucleic acid (synthetic DNA)
CC SEQUENCE 60 BP; 16 A; 13 C; 19 G; 12 T; 0 OTHER.
SQ

Query Match          50.6%; Score 160; DB 4; Length 60;
Best Local Similarity 62.2%; Pred. No. 4.20e-04;
Matches 23; Conservative 9; Mismatches 5; Indels 0; Gaps 0;

Db      24 GAGGAGTCTTGACCCCAAGAACAATTGAGTGTTGG 60
Oy      1 GARGARTTYVTMACNCNARARAARYTCARTGYGTNG 37
||:::||::| | | | |::||:| | | | |
|||||

RESULT 15
ID US-08-940-391-1 STANDARD; DNA; UNC; 2B15 BP.
AC xxxxxx
DT
DE Sequence 1, Application US/08940391
CC Sequence 1, Application US/08940391
CC Patent No. 5965373
CC GENERAL INFORMATION:
CC APPLICANT: Zimmermann, Rainer; Park, John E.;
CC APPLICANT: Rettig, Wolfgang; Old, Lloyd J.
CC TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION
CC NUMBER OF SEQUENCES: 10
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Felte & Lynch
CC STREET: 805 Third Avenue
CC CITY: New York City
CC STATE: New York
CC COUNTRY: USA
CC ZIP: 10022
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage
CC COMPUTER: IBM PS/2
CC OPERATING SYSTEM: PC-DOS
CC SOFTWARE: Wordperfect
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/940,391
CC FILING DATE: 01-OCT-1997
CC CLASSIFICATION: 530
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/619,280
CC FILING DATE: 18-MARCH-1996
CC APPLICATION NUMBER: 08/230,491
CC FILING DATE: 20-APRIL-1994
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Hanson, No. 5965373man D.
CC REGISTRATION NUMBER: 30,946
CC REFERENCE/DOCKET NUMBER: LUD 5530.1
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 688-9200
CC TELEFAX: (212) 838-3884
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
```



```

ACCESSIONS
22-Jun-1999
A32297; A32423; S03604; S05468; S05467; A32546; S02239;
A26557; C31567; S00232; A23937; G01351
REFERENCE
#authors
Riegman, P.H.J.; Vlietstra, R.J.; van der Korput, J.A.G.M.;
#journal
Romijn, J.C.; Tjepman, J.
#title
Biochem. Biophys. Res. Commun. (1989) 159:95-102
Characterization of the prostate-specific antigen gene: a
novel human kallikrein-like gene.
#cross-references MUID:89165891
#accession
A32297
#status
preliminary
##molecule-type DNA
##residues
1-261 #label R1E
REFERENCE
#authors
Lundwall, A.
#journal
Biochem. Biophys. Res. Commun. (1989) 161:1151-1159
#title
Characterization of the gene for prostate-specific antigen, a
human glandular kallikrein.
#cross-references MUID:89302090
#accession
A32423
#status
preliminary
##molecule-type DNA
##residues
1-261 #label LUN
#cross-references GB:M27274; NID:g190552; PIDN:AAA60192.4; PID:g190553
REFERENCE
#authors
Digby, M.; Zhang, X.Y.; Richards, R.I.
#journal
Nucleic Acids Res. (1989) 17:2137
#title
Human prostate specific antigen (PSA) gene: structure and
linkage to the kallikrein-like gene, hKc-1.
#cross-references MUID:89183632
#accession
S03604
##molecule-type DNA
##residues
1-261 #label DIG
#cross-references EMBL:X13940
REFERENCE
#authors
Klobeck, H.G.; Combrinato, G.; Schulz, P.; Arbusow, V.;
Fittler, F.
#submission
submitted to the EMBL Data Library, May 1989
#accession
S05468
##molecule-type DNA
##residues
1-261 #label KLI
#cross-references EMBL:X14810; NID:g35732; PIDN:CAA32915.1; PID:g296671
REFERENCE
#authors
Klobeck, H.G.; Combrinato, G.; Schulz, P.; Arbusow, V.;
Fittler, F.
#journal
Nucleic Acids Res. (1989) 17:3981
#title
Genomic sequence of human prostate specific antigen (PSA).
#cross-references MUID:89282407
#accession
S05467
##molecule-type DNA
##residues
1-29 #label KL2
#cross-references EMBL:X14810
REFERENCE
#authors
Henttu, P.; Vinko, P.
#journal
Biochem. Biophys. Res. Commun. (1989) 160:903-910
#title
cDNA coding for the entire human prostate specific antigen
shows high homologies to the human tissue kallikrein genes.
#cross-references MUID:89246551
#accession
A32546
##molecule-type RNA
##residues
1-72 'T', 74-85 'I', 87-174, 'P', 176-183, 'O', 185-259, 'D',
261 #label HEN
#cross-references GB:M26663
REFERENCE
#authors
Schulz, P.; Stucka, R.; Feldmann, H.; Combrinato, G.; Klobeck,
H.G.; Fittler, F.
#journal
Nucleic Acids Res. (1989) 16:6226
#title
Sequence of a cDNA clone encompassing the complete mature
human prostate specific antigen (PSA) and an unspliced
leader sequence.
#cross-references MUID:86289366

```

```

#accession      S02239
##molecule_type mRNA
##residues      17-63,'T','G','A','C','G','T','A','G','M'
##cross-references EMBL:X07730
REFERENCE
#authors         Lundvall, A.; Lilja, H.
#journal         FEBS Lett. (1987) 214:317-322
#title           Molecular cloning of human prostate specific antigen cDNA.
#cross-references MIMD:87190978
#accession       A26757
##molecule_type mRNA
##residues       5-261 ##label L02
REFERENCE
#authors         Rieeman, P.H.J.; Klaassen, P.; van der Korput, J.A.G.M.;
                Romijn, J.C.; Trapman, J.
#journal         Biochem. Biophys. Res. Commun. (1988) 155:181-188
#title           Molecular cloning and characterization of novel prostate
                antigen cDNA's.
#cross-references MIMD:88326297
#accession       C31567
##molecule_type mRNA
##residues       5-261 ##label RT2
##cross-references GB:M21895; NID:g189523; PIDN:AAA59995.1; PID:g189522
                the authors translated the codon GGC for residue 28 as
                Arg and TGG for residue 29 as Leu
#note            S00232
REFERENCE
#authors         Schaller, J.; Akiyama, K.; Tsuda, R.; Hara, M.; Mattl, T.;
                Rickli, E.E.
#journal         Eur. J. Biochem. (1987) 170:111-120
#title           Isolation, characterization and amino-acid sequence of
                gamma-seminoprotein, a glycoprotein from human seminal
                plasma.
#cross-references MIMD:88082806
#accession       S00232
##molecule_type protein
##residues       25-261 ##label SC2
REFERENCE
#authors         Watt, K.W.K.; Lee, P.J.; W'Tinkulu, T.; Chan, W.P.; Loor, R.
#journal         Proc. Natl. Acad. Sci. U.S.A. (1986) 83:3165-3170
#title           Human prostate-specific antigen: structural and functional
                similarity with serine proteases.
#cross-references MIMD:86205857
#accession       A23937
##molecule_type protein
##residues       25-93,'T','G','A','C','G','T','A','G','M','I','V','Q','R','S','K'
                ##label MAT
REFERENCE
#authors         Moreno, J.M.
#submission      Submitted to the EMBL Data Library, November 1994
#accession       G01551
##status          translated from GB/EMBL/DDBJ
##molecule_type mRNA
##residues        1-261 ##label MOR
##cross-references EMBL:U17040; NID:g595945; PIDN:AAA56764.1;
                PID:g595946
GENETICS
#gene            GDB:APS
##cross-references GDB:119695; OMIM:176820
#map_position     19q13.3-19q13.3
#introns          16/1; 69/2; 165/1; 210/3
CLASSIFICATION    #superfamily trypsin; trypsin homology
KEYWORDS          glycoprotein; hydrolase; prostate; serine proteinase
FEATURES
1-17              #domain signal sequence #status predicted #label SIG\
18-24             #domain propeptide #status predicted #label PRO\
25-261            #product prostatic serine proteinase #status
                experimental #label MAT\
25-253            #domain trypsin homology #label TRY\
65,120,213        #active site His, Asp, Ser #status predicted
SUMMARY           #length 261 #molecular-weight 28741 #checksum 1256

Query Match      100.0%; Score 183; DB 2; Length 261;
Best Local Similarity 100.0%; Pred. No. 1,266-2;

```

(TM)

Distribution rights by oxidized molecular species

protein - protein database search/ using omics .....?

Thu Jun 22 21:44:37 2000, 203.580 Million cell updates/sec

generalized

(1-25) from US09332866.pdf

1 EEF1P KKLQCVDLHVISNDVCAQV 23

PAM 150  
Can 15

142080 seqs, 4/1/2406 residues

Listing first 45 summaries

1:pir1 2:pir2 3:pir3 4:pir4

Mean 28.844; Variance 40.415; Scale 0.114

ived by analysis of the total score distribution.

## SUMMARIES

Accession	Length	DB	ID	Description	Pred. No.
U00000.0	234	2	A3167	prostatic proantigen	1,266-28
U00000.0	261	2	A3327	prostate-specific ser	1,266-28
U00001.3	261	2	A32711	tissue kallikrein (EC	1,946-248
U00001.0	261	2	A29586	tissue kallikrein (EC	4,466-153
U00001.0	261	2	A45930	tissue kallikrein (EC	8,226-10
U00001.0	261	2	A5303	tissue kallikrein (EC	1,406-00
U00001.0	189	2	I52712	gene PSA protein - hu	4,476-00
U00001.0	257	2	S33772	tissue kallikrein (EC	5,668-00
U00001.0	262	1	KOH	tissue kallikrein (EC	5,668-00
U00001.0	263	2	A15686	tissue kallikrein (EC	5,686-00
U00001.0	260	2	A37938	prostatic arginine es	1,556-00
U00001.0	260	2	S01971	tissue kallikrein (EC	1,556-00
U00001.0	261	2	S1565	tissue kallikrein, glandular	1,816-00
U00001.0	250	2	A29745	tissue kallikrein (EC	2,946-00
U00001.0	244	2	A44284	tissue kallikrein (EC	4,766-00
U00001.0	259	1	KORFN	tonin (EC 3.4.21.-) p	7,696-00
U00001.0	156	2	B32863	tissue kallikrein (EC	7,696-00
U00001.0	188	2	B32300	tissue kallikrein (EC	7,696-00
U00001.0	261	2	A31136	tissue kallikrein (EC	7,696-00
U00001.0	261	2	A34079	tissue kallikrein (EC	1,996-00
U00001.0	261	1	KOMSI	tissue kallikrein (EC	3,186-00
U00001.0	94	2	PC2013	tissue kallikrein (EC	3,186-00
U00001.0	246	2	J01471	trypsin (EC 3.4.21.4)	3,186-00
U00001.0	246	2	J01472	trypsin (EC 3.4.21.4)	3,186-00

24	78	259	B3136	tissue kallikrein (EC	3.18e-03
25	78	251	A25605	tissue kallikrein (EC	3.18e-03
26	77	259	D23863	tissue kallikrein (EC	5.08e-03
27	76	96	A03308	tissue kallikrein (EC	8.08e-03
28	76	256	NGMS4	7S nerve growth factor	8.08e-03
29	75	259	A29746	tissue kallikrein (EC	1.28e-02
30	72	81	A18966	tissue kallikrein (EC	5.01e-02
31	72	241	S59048	tissue kallikrein (EC	5.01e-02
32	71	237	KOP3	tissue kallikrein (EC	7.85e-02
33	71	244	S72219	chymotrypsin B - Atla	1.90e-01
34	69	240	S59047	trypsin (EC 3.4.21.4)	1.90e-01
35	69	265	KORIP	tissue kallikrein (EC	2.94e-01
36	68	239	A27207	tissue kallikrein (EC	2.94e-01
37	68	259	S68424	allergen Der f III pr	2.94e-01
38	68	343	A57014	proelastin (EC 3.4.21.4)	4.53e-01
39	67	250	T01779	trypsin (EC 3.4.21.4)	4.53e-01
40	67	250	S31384	trypsin (EC 3.4.21.4)	4.53e-01
41	67	253	A23968	serine proteinase SCC	4.53e-01
42	67	261	A41020	tissue kallikrein (EC	4.53e-01
43	67	261	J50236	tissue kallikrein (EC	4.53e-01
44	67	616	A20304	prolactin receptor 2	4.53e-01
45	66	238	S31779	trypsin (EC 3.4.21.4)	6.96e-01

## ALIGNMENTS

RESULT	1
ENTRY	A31567
TITLE	#type fragment prostatic p10antigen precursor (clone 525) - human (fragment,
ORGANISM	#formal name Homo sapiens #common_name man
DATE	31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 22-Jun-1999

ACCESSIONS  
A31567  
REFERENCE  
A90144

P u t : Klaassen, P.: van der Korpuit, J.A.G.M.

Romijn, J.C.; Trapman, J.  
Biochem. Biophys. Res. Commun. (1988) 155:181-188

**#Title** Molecular cloning and characterization of novel prostate cancer-specific cDNAs  
**#References** WITD.88326297

```
#molecule_type mRNA
#accession A31567
#cross_references MolDb, UniProt
```

```

#residues      1-234  #label RIE
#Cross-references GB:M21896; NID:g189525; P1TN:AAA59996.1; PID:g189526
#...the authors translated the codon GGC for residue 24 as

```

CLASSIFICATION #superfamily trypsin; trypsin homology  
#hole Arg and TGG for residue 25 as Leu  
One unusual characteristic

```
FEATURE
1-13      #domain signal sequence (fragment)  #status predicted
          #label sig\
```

	domain	trypsin	homology	#status	predicted	#label	PRO
14-21	#domain	trypsin	homology	#status	predicted	#label	PRO
21-234	#domain	trypsin	homology	#status	atypical	#label	TRY

22-234	#product	prostatic antigen	#status	predicted	#label	wat
61,116,209	#region	defective catalytic triad				
	#checksum	806				

Query Match	100.0%;	Score 183;	DB 2;	Length 234;
SUMMARY	#LENGTH 234	#CHARCOUNT 234		

```

Best Local Similarity 100.0%; Pred. No. 1.26e-28;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Db 159 EEFLTPKKLQCYDLHVISNDYCAQV 183

QY 1 EEFLPKKLQCVDLHVISNDYCAQV 25

RESULT	2
ENTRY	A32297
	#type complete

TITLE	PROSTATE-SPECIFIC SERINE PROTEINASE (EC 3.4.21.-) PRECURSOR HUMAN
1	prostate-specific serine proteinase (EC 3.4.21.-) precursor human

ORGANISM	
#formal name	Homo sapiens #Common_name man
ALTERNATE_NAMES	gamma-semiophoretin; prostatec antigen; prostatic 3.4.21.35)-related proteinase, prostatic

```

08-Sep-1989 #sequence_revision 08-Sep-1989 #text_change
DATE

```

Matches 25: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 163 EEFLLPKKLCQVDLHVISNDVCAQV 187  
1 EEFLLPKKLCQVDLHVISNDVCAQV 25

RESULT 3  
ENTRY 3  
TITLE S35711 #type complete  
tissue kallikrein (EC 3.4.21.35) precursor, prostatic -  
rhesus macaque

ALTERNATE\_NAMES prostate-specific antigen; prostate-specific serine  
protease

ORGANISM #formal\_name Macaca mulatta #common\_name rhesus macaque  
06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change  
22-Jun-1999

DATE S35711: S34239

ACCESSIONS S35711  
REFERENCE S35711

#authors Gauthier, E.R.; Chappelaire, P.; Tremblay, R.R.; Dube, J.Y.  
#journal Biochim. Biophys. Acta (1993) 1174:207-210  
#title Characterization of rhesus monkey prostate specific antigen  
cDNA.

#cross-references MUID:93363642

#accession S35711

#molecule-type mRNA

#residues 1-261 ##label GAU

#cross-references EMBL:X73560; NID:9311843; PIDN:CA51957.1;  
PID:9311844

CLASSIFICATION #superfamily trypsin; trypsin homology  
KEYWORDS glycoprotein; hydrolase; prostate; serine proteinase  
FEATURE 1-24

24-261 #domain signal sequence #status predicted #label SIG  
25-253 #product kallikrein #status predicted #label MAT  
65,120,213 #domain trypsin homology #label TRY  
#active-site His, Asp, Ser #status predicted

SUMMARY #length 261 #molecular-weight 28816 #checksum 3070

Query Match  
Best Local Similarity 92.0%; Score 167; DB 2; Length 261;  
Matches 23; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 163 EEFLLPKKLCQVDLHVISNDVCAQV 187  
1 EEFLLPKKLCQVDLHVISNDVCAQV 25

RESULT 4  
ENTRY 4  
TITLE A29586 #type complete  
tissue kallikrein (EC 3.4.21.35) hGK-1 precursor - human  
glandular kallikrein

ALTERNATE\_NAMES #formal\_name Homo sapiens #common\_name man  
08-Mar-1989 #sequence\_revision 08-Mar-1989 #text\_change  
22-Jun-1999

DATE A29586

ACCESSIONS A29586  
REFERENCE A29586

#authors Schedlich, L.J.; Bennetts, B.H.; Morris, B.J.  
#journal DNA (1987) 6:429-437  
#title Primary structure of a human glandular kallikrein gene.  
#cross-references MUID:88054467

#accession A29586

#molecule-type DNA

#residues 1-261 ##label SCH

#cross-references GB:M18157; NID:9186640; PIDN:AAA74454.1; PID:9386842

#note The authors translated the codon TAC for residue 43 as  
Trp

GENETICS #introns 16/1: 69/2, 165/1: 210/3  
CLASSIFICATION #superfamily trypsin; trypsin homology  
KEYWORDS hydrolase; serine proteinase  
FEATURE 23-253

23-253 #domain trypsin homology #label TRY  
65,140,213 #active-site His, Asp, Ser #status predicted

SUMMARY #length 261 #molecular-weight 28671 #checksum 807

Query Match  
Best Local Similarity 71.0%; Score 130; DB 2; Length 261;  
Matches 16; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Db 163 EEFLLPKKLCQVDLHVISNDVCAQV 186  
1 EEFLLPKKLCQVDLHVISNDVCAQV 24

RESULT 5  
ENTRY 5  
TITLE S45303 #type complete  
tissue kallikrein (EC 3.4.21.35) precursor - dog

ALTERNATE\_NAMES #formal\_name Canis lupus familiaris #common\_name dog  
20-Oct-1994 #sequence\_revision 03-Aug-1995 #text\_change  
22-Jun-1999

DATE S45303: S38487

ACCESSIONS S45303  
REFERENCE S45303

#authors Gauthier, E.R.; Dumas, C.; Chappelaire, P.; Tremblay, R.R.;  
Dube, J.Y.  
#journal Biochim. Biophys. Acta (1994) 1218:102-104  
#title Characterization of canine pancreas kallikrein cDNA.  
#cross-references MUID:94250683

#accession S45303

#molecule-type mRNA

#residues 1-261 ##label GAU

#cross-references EMBL:X75479; NID:9414018; PIDN:CA53210.1;  
PID:9414019

CLASSIFICATION #superfamily trypsin; trypsin homology  
KEYWORDS hydrolase; serine proteinase  
FEATURE 1-24

25-261 #domain signal sequence #status predicted #label SIG  
25-253 #product tissue kallikrein #status predicted #label MAT  
65,120,213 #domain trypsin homology #label TRY  
#active-site His, Asp, Ser #status predicted

SUMMARY #length 261 #molecular-weight 28913 #checksum 6470

Query Match  
Best Local Similarity 59.0%; Score 108; DB 2; Length 261;  
Matches 13; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Db 163 DFFLPDDLCQVDLELLSDICA 185  
1 EEFLLPKKLCQVDLHVISNDVCA 23

RESULT 6  
ENTRY 6  
TITLE I52712 #type fragment  
gene PSA protein - human (fragment)

ALTERNATE\_NAMES #formal\_name Homo sapiens #common\_name man  
01-Nov-1996 #sequence\_revision 01-Nov-1996 #text\_change  
18-Jul-1997

DATE I52712

ACCESSIONS I52712  
REFERENCE I52712

#authors Wonne, M.; Croce, C.M.; Yu, H.; Diamandis, E.P.  
#journal Cancer Res. (1994) 54:6344-6347  
#title Molecular characterization of prostate-specific antigen  
messenger RNA expressed in breast tumors.  
#cross-references MUID:95079406

#accession I52712

#status preliminary; translated from GB/EMBL/DBJ

#molecule-type mRNA

#residues 1-189 ##label RES

#cross-references GB:S75755; NID:9681469

GENETICS #gene PSA  
CLASSIFICATION #superfamily trypsin; trypsin homology  
FEATURE 38-189

38-189 #domain trypsin homology (fragment) #label TRY  
SUMMARY #length 189 #checksum 5273

Query Match  
Best Local Similarity 100.0%; Score 107; DB 2; Length 189;  
Matches 100; Conservative 100; Mismatches 0; Indels 0; Gaps 0;

Matches 14: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Db 176 EEFLLPKKLCQVDL 189  
 |||  
 1 EEFLLPKKLCQVDL 14

QY

RESULT 7  
 ENTRY S33772 #type complete  
 TITLE tissue kallikrein (EC 3.4.21.35) precursor, renal -  
 #formal name Macaca fascicularis #common\_name crab-eating  
 macaque  
 ORGANISM 22-Nov-1993 #sequence\_revision 26-May-1995 #text\_change  
 22-Jun-1999

ACCESSIONS  
 REFERENCE S33772  
 #authors Lin, F.K.; Lin, C.H.; Chou, C.C.; Chen, K.; Lu, H.S.;  
 Bacheller, W.; Herrera, C.; Jones, T.; Chao, J.; Chao, L.  
 #journal Biochim. Biophys. Acta (1993) 1173:325-328  
 #title Molecular cloning and sequence analysis of the monkey and  
 human tissue kallikrein genes.  
 #cross-references MUID:93305727  
 #accession S33772  
 #status preliminary  
 #molecule\_type mRNA  
 #residues 1-257 #label LIN  
 #cross-references GB:J10039; NID:q293140; PIDN:AAA36853.1; PID:q293141  
 #note the authors translated the codon CGT for residue 17 as  
 Ala, GCG for residue 18 as Arg, GCG for residue 54 as  
 Arg, TCG for residue 71 as Met, and AAT for residue  
 160 as His

CLASSIFICATION #superfamily trypsin; trypsin homology  
 KEYWORDS hydrolyase; serine proteinase

FEATURES  
 25-249 #domain trypsin homology #label TRY  
 62,116,209 #active-site His, Asp, Ser #status predicted  
 #length 257 #molecular\_weight 28237 #checksum 7941

SUMMARY

Query Match 52.5%; Score 96; DB 2; Length 257;  
 Best Local Similarity 56.5%; Pred. No. 4.47e-07;  
 Matches 13; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Db 159 ENFSFPDLCVDEILPNDCEA 181  
 |||  
 1 EEFLLPKKLCQVDLHVISNDVCA 23

QY

RESULT 8  
 ENTRY KOHU #type complete  
 TITLE tissue kallikrein (EC 3.4.21.35) precursor - human  
 ALTERNATE\_NAMES kininogenin; pancreatic kallikrein; salivary kallikrein;  
 urinary kallikrein  
 ORGANISM #formal name Homo sapiens #common\_name man  
 DATE 30-Sep-1987 #sequence\_revision 30-Sep-1987 #text\_change  
 18-Jun-1999

ACCESSIONS  
 REFERENCE A24696; A28678; A23587; S05642; JX0040; A04628; A60248;  
 JK0218; A60251; A60250; S33773; S55240; S55241; A27256;  
 A60247; J70010; S03367  
 A24696  
 #authors Fukushima, D.; Kitamura, N.; Nakanishi, S.  
 #journal Biochemistry (1985) 24:8037-8043  
 #title Nucleotide sequence of cloned cDNA for human pancreatic  
 kallikrein.  
 #cross-references MUID:86131629  
 #accession A24696  
 #molecule\_type mRNA  
 #residues 1-262 #label FUK  
 #cross-references GB:M25629; NID:q186652; PIDN:AAA36136.1; PID:q186653

REFERENCE  
 #authors Evans, B.A.; Yun, Z.X.; Close, J.A.; Tregear, G.W.; Kitamura,  
 N.; Nakanishi, S.; Callen, D.F.; Baker, E.; Hyland, V.J.;  
 Sutherland, G.R.; Richards, R.I.

#journal Biochemistry (1988) 27:3124-3129  
 #title Structure and chromosomal localization of the human renal  
 kallikrein gene.  
 #cross-references MUID:88269498  
 #accession A28678  
 #molecule\_type DNA  
 #residues 1-185, 'K', 187-262 #label EVA  
 #cross-references GB:M33105; GB:M33106; GB:M33107; GB:M33108;  
 GB:M33109; NID:q186649; PIDN:AAA59455.1; PID:q186651

REFERENCE  
 #authors Baker, A.R.; Shine, J.  
 #journal DNA (1985) 4:445-450  
 #title Human kidney kallikrein: cDNA cloning and sequence analysis.  
 #cross-references MUID:86155264  
 #accession A23587  
 #molecule\_type mRNA  
 #residues 17-262 #label BAK

REFERENCE  
 #authors Angermann, A.; Bergmann, C.; Appelhans, H.  
 #journal Blochem. J. (1989) 262:787-793  
 #title Cloning and expression of human salivary gland kallikrein in  
 Escherichia coli.  
 #cross-references MUID:90073574  
 #accession S05642  
 #molecule\_type mRNA  
 #residues 1-144, 'E', 146-185, 'K', 187-262 #label ANG  
 #cross-references EMBL:X13561; NID:q34026; PIDN:CAA31912.1; PID:q34027  
 #note the authors translated the codon GGT for residue 226 as  
 Glu

REFERENCE  
 #authors Takahashi, S.; Irie, A.; Miyake, Y.  
 #journal J. Biochem. (1988) 104:22-29  
 #title Primary structure of human urinary prokallikrein.  
 #cross-references MUID:89125217  
 #accession JX0040  
 #molecule\_type protein  
 #residues 18-262 #label TAK

REFERENCE  
 #authors Lotzspeich, F.; Geiger, R.; Henschen, A.; Kutzbach, C.  
 #journal Hoppe-Seyler's Z. Physiol. Chem. (1979) 360:1947-1950  
 #title N-terminal amino acid sequence of human urinary kallikrein  
 homology with other serine proteases.  
 #cross-references MUID:80114126  
 #accession A04628  
 #molecule\_type protein  
 #residues 25-55 #label LOT

REFERENCE  
 #authors Kellermann, J.; Lotzspeich, F.; Geiger, R.; Deutzmann, R.  
 #journal Adv. Exp. Med. Biol. (1989) 247A:519-525  
 #title Human urinary kallikrein: amino acid sequence and  
 carbohydrate attachment sites.  
 #cross-references MUID:90102100  
 #accession A60248  
 #molecule\_type protein  
 #residues 25-262 #label KEL

REFERENCE  
 #authors Lu, H.S.; Lin, F.K.; Chao, L.; Chao, J.  
 #journal Int. J. Pept. Protein Res. (1989) 33:237-249  
 #title Human urinary kallikrein: complete amino acid sequence and  
 sites of glycosylation.  
 #cross-references MUID:89326688  
 #accession JK0218  
 #molecule\_type protein  
 #residues 25-185, 'K', 187-262 #label LUH

REFERENCE  
 #authors Terashima, H.; Atomi, Y.; Ohnishi, N.; Kuroda, A.; Morioka,  
 Y.; Ikeita, M.; Aoki, K.; Kamada, M.; Kituki, K.; Moriya,  
 H.  
 #journal Adv. Exp. Med. Biol. (1989) 247B:177-182  
 #title Characterization of human pancreatic kallikrein.  
 #cross-references MUID:90119035  
 #accession A60251



```

Best Local Similarity 50.0%; Pred.No. 1.55e-05;
Matches 10; conservative 5; Mismatches 5; Indels 0; Gaps 0;

Db 165 FQTPDDLCQVSIKLPNEVC 184
      1 11 1111 : ::1:11
QY 3 FLTPKKIQCVDLHVISNDYCA 22

RESULT 12
ENTRY S15685 #type fragment
TITLE kallikrein, glandular - multimammate rat (Mastomys natalensis) (fragment)
ORGANISM #formal_name Mastomys natalensis
DATE 09-Jun-1994 #sequence_revision 10-Nov-1995 #text_change 22-Jun-1999

ACCESSIONS S15685
REFERENCE S15685
AUTHORS Bell, R.A.; Fahnestock, M.
SUBMISSION submitted to the EMBL Data Library, December 1989
DESCRIPTION Sequence of a cDNA clone coding for a glandular kallikrein from Mastomys natalensis.
#accession S15685
#status preliminary
#molecule_type mRNA
#residues 1-250 ##label BEL
#cross-references EMBL:X17351; NID:g55524; PIDN:CA435231.1; PID:e32221;
PID:g1334162
CLASSIFICATION #superfamily trypsin; trypsin homology
FEATURE #domain trypsin homology #label TRY\
22-250 #active_site His, Asp, Ser #status predicted
62,117,211 #length 250 #checksum 4256
SUMMARY

Query Match 47.5%; Score 87; DB 2; Length 250;
Best Local Similarity 47.8%; Pred. No. 4.17e-05;
Matches 11; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Db 161 EEYKFPDDLCQVYLKIMPNEICA 183
      11 : 1 1111 : ::1:11
QY 1 EEFLTPKKIQCVDLHVISNDYCA 23

RESULT 13
ENTRY A29745 #type complete
TITLE tissue kallikrein (EC 3.4.21.35) mGK-9 precursor, submandibular - mouse
ALTERNATE_NAMES glandular kallikrein mGK-9; major epidermal growth factor-binding protein
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 22-Jun-1999
ACCESSIONS C29746; A29745; A27120; I70015
REFERENCE A90522
AUTHORS Drinkwater, C.C.; Evans, B.A.; Richards, R.I.
JOURNAL Biochemistry (1987) 26:6750-6756
TITLE Mouse glandular kallikrein genes: identification and characterization of the genes encoding the epidermal growth factor binding proteins.

#cross-references MUID:88107594
#accession C29746
#molecule_type DNA
#residues 1-261 ##label BRI
#cross-references GB:M17985; NID:g193476; PIDN:AA37681.1; PID:g387166
#experimental_source strain BALB/c; salivary gland
REFERENCE A29745
AUTHORS Blaber, M.; Isaacson, P.J.; Bradshaw, R.A.
JOURNAL Biochemistry (1987) 26:6742-6749
TITLE A complete cDNA sequence for the major epidermal growth factor binding protein in the male mouse submandibular gland.
#cross-references MUID:88107593
#accession A29745
#molecule_type mRNA

```



```

#journal      J. Biochem. (1987) 102:1389-1404
#title        Characterization of serine proteinases isolated from rat
              submaxillary gland: with special reference to the
              degradation of rat kininogens by these enzymes.
#cross-references MUID:8819805/
#accession    C41429
#status       preliminary
#molecule_type protein
#residues     10-61 #label KAT
CLASSIFICATION #superfamily trypsin; trypsin homology
KEYWORDS       hydrolase; serine proteinase
FEATURE        10-236
              48-103,196
SUMMARY        #domain trypsin homology #label TRY\
              #active_site His, Asp, Ser #status predicted
              #length 244 #checksums 6745

Query Match      45.4% Score 83: DB 2: Length 244;
Best local Similarity 47.4%; Pred No 2,94e-04;
Matches 9; Conservative 6; Mismatches 4; Indels 0; Gaps 0

Db 151 PDDLQCVNHLSENKCE 169
      | | | | | : | | | | : | :
QY 6 PKRLQCVDLHVISNDVCAQ 24

RESULT 15
ENTRY KORTN #type complete
TITLE tonin (EC 3.4.4.1.) precursor - rat
ALTERNATE_NAMES esterase 1
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE 25-Feb-1985 #sequence_revision 02-Dec-1994 #text_change
      18-Jun-1999
ACCESSIONS B33359; A33340; C23863; A93323; A94635; A34050; A30971;
A00945

REFERENCE
#authors      Wines, D.R.; Brady, J.M.; Pritchett, D.B.; Roberts, J.L.;
              MacDonald, R.J.
#journal      J. Biol. Chem. (1989) 264:7653-7662
#title        Organization and expression of the rat kallikrein gene
              family.
#cross-references MUID:89214217
#accession    B33359
#molecule_type DNA
#residues     1-259 #label WIN
#cross-references GB:M23877; GB:O04702; GB:M23878; NID:g207411;
              PIDN:AAA42259.1; PID:g207413; GB:M23879
              PIDN:AAA42259.1; PID:g207413; GB:M23879

REFERENCE
#authors      Sha1, S.Y.; Woodley-Miller, C.; Chao, J.; Chao, L.
#journal      Biochemistry (1989) 28:5334-5343
#title        Characterization of genes encoding rat tonin and a
              kallikrein-like serine protease.
#cross-references MUID:89375248
#accession    A33340
#molecule_type DNA
#residues     1-259 #label SHA
#cross-references GB:M25533; NID:g206775; PIDN:AAA42081.1; PID:g206776;
              GB:U02860

REFERENCE
#authors      Ashley, P.L.; MacDonald, R.J.
#journal      Biochemistry (1985) 24:4512-4520
#title        Kallikrein-related mRNAs of the rat submaxillary gland:
              nucleotide sequences of four distinct types including
              tonin.
#cross-references MUID:86051477
#accession    C23863
#molecule_type mRNA
#residues     1-259 #label ASH
#cross-references GB:M11565; NID:g205033; PIDN:AAA41466.1; PID:g205034
              PIDN:AAA41466.1; PID:g205034

REFERENCE
#authors      Lazure, C.; Leduc, R.; Seidah, N.G.; Thibault, G.; Genest,
              J.; Chretien, M.
#journal      Nature (1984) 307:555-558

```

```

#journal      J. Biochem. (1987) 102:1389-1404
#title        Characterization of serine proteinases isolated from rat
              submaxillary gland: with special reference to the
              degradation of rat kininogens by these enzymes.
#cross-references MUID:8819805/
#accession    C41429
#status       preliminary
#molecule_type protein
#residues     10-61 #label KAT
CLASSIFICATION #superfamily trypsin; trypsin homology
KEYWORDS       hydrolase; serine proteinase
FEATURE        10-236
              48-103,196
SUMMARY        #domain trypsin homology #label TRY\
              #active_site His, Asp, Ser #status predicted
              #length 244 #checksums 6745

Query Match      45.4% Score 83: DB 2: Length 244;
Best local Similarity 47.4% Pred No 2,94e-04;
Matches 9; Conservative 6; Mismatches 4; Indels 0; Gaps 0

Db 151 PDDLQCVNHLSENKCE 169
      | | | | | : | | | | : | :
QY      6 PKRLQCVDLHVISNDVCAQ 24

RESULT 15
ENTRY   KORTIN
TITLE   tonin (EC 3.4.4.1.) precursor - rat
ALTERNATE_NAMES esterase 1
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE     25-Feb-1985 #sequence_revision 02-Dec-1994 #text_change
              18-Jun-1999
ACCESSION B33359; A33340; C23863; A93323; A94635; A34050; A30971;
A00945

REFERENCE
#authors     Wines, D.R.; Brady, J.M.; Pritchett, D.B.; Roberts, J.L.;
              MacDonald, R.J.
#journal     J. Biol. Chem. (1989) 264:7653-7662
#title       Organization and expression of the rat kallikrein gene
              family.
#cross-references MUID:89214217
#accession   B33359
#molecule_type DNA
#residues    1-259 #label WIN
#cross-references GB:M23877; GB:O04702; GB:M23878; NID:g207411;
              PIDN:AAA42259.1; PID:g207413; GB:M23879
              PIDN:AAA42259.1; PID:g207413; GB:M23879

REFERENCE
#authors     Sha1, S.Y.; Woodley-Miller, C.; Chao, J.; Chao, L.
#journal     Biochemistry (1989) 28:5334-5343
#title       Characterization of genes encoding rat tonin and a
              kallikrein-like serine protease.
#cross-references MUID:89375248
#accession   A33340
#molecule_type DNA
#residues    1-259 #label SHA
#cross-references GB:M25533; NID:g206775; PIDN:AAA42081.1; PID:g206776;
              GB:U02860

REFERENCE
#authors     Ashley, P.L.; MacDonald, R.J.
#journal     Biochemistry (1985) 24:4512-4520
#title       Kallikrein-related mRNAs of the rat submaxillary gland:
              nucleotide sequences of four distinct types including
              tonin.
#cross-references MUID:86051477
#accession   C23863
#molecule_type mRNA
#residues    1-259 #label ASH
#cross-references GB:M11565; NID:g205033; PIDN:AAA41466.1; PID:g205034
              PIDN:AAA41466.1; PID:g205034

REFERENCE
#authors     Lazure, C.; Leduc, R.; Seidah, N.G.; Thibault, G.; Genest,
              J.; Chretien, M.
#journal     Nature (1984) 307:555-558

```



\*\*\*\*\*  
 WIRE (TM)  
 \*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

MSPrch\_bp protein - protein database search, using Smith-Waterman algorithm  
 Run on: Thu Jun 22 21:22:03 2000; MasPar time 3.84 Seconds  
 Tabular output not generated. 198.438 Million cell updates/sec

Title: >US-09-332-866-1  
 Description: (1-25) from US09332866.pep  
 Perfect Score: 183  
 Sequence: 1 EEFLTPKIKLQCVLHVISNDVCAQV 25

Scoring table: PAM 150  
 Gap 15

Searched: 83857 seqs, 30454973 residues

Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database: swiss-prot38  
 1:swissprot

Statistics: Mean 29.602; Variance 37.132; scale 0.797

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	183	100.0	261	1	PROS_HUMAN	PROSTATE SPECIFIC ANTI	4.77e-32
2	167	91.3	261	1	PROS_MACMU	PROSTATE SPECIFIC ANTI	2.12e-27
3	130	71.0	261	1	KLK2_HUMAN	GLANDULAR KALLIKREIN 2	5.07e-17
4	96	52.5	257	1	KLK1_MACFA	GLANDULAR KALLIKREIN 1	5.50e-08
5	95	51.9	258	1	KLK1_PAPHA	GLANDULAR KALLIKREIN 1	6.15e-08
6	91	49.7	262	1	KLK1_HUMAN	GLANDULAR KALLIKREIN 1	5.75e-07
7	91	49.7	263	1	KLK1_PAPHA	GLANDULAR KALLIKREIN 1	5.75e-07
8	89	48.6	260	1	ESTR_CANFA	ANGIOTENSIN ESTERASE PREC	1.73e-06
9	89	48.6	261	1	KLK4_MOUSE	GLANDULAR KALLIKREIN 4	1.73e-06
10	84	45.9	261	1	KLK3_MOUSE	GLANDULAR KALLIKREIN 3	2.59e-05
11	83	45.4	264	1	KLK3_MOUSE	GLANDULAR KALLIKREIN 3	2.59e-05
12	82	44.8	259	1	KLK3_MOUSE	GLANDULAR KALLIKREIN 3	2.59e-05
13	81	44.3	188	1	KLK3_MOUSE	GLANDULAR KALLIKREIN 3	2.59e-05
14	81	44.3	251	1	KLK3_MOUSE	GLANDULAR KALLIKREIN 3	2.59e-05
15	81	44.3	251	1	KLK3_MOUSE	GLANDULAR KALLIKREIN 3	2.59e-05
16	79	43.2	261	1	KLK3_MOUSE	GLANDULAR KALLIKREIN 3	2.59e-05
17	78	42.6	246	1	KLK3_MOUSE	GLANDULAR KALLIKREIN 3	2.59e-05
18	78	42.6	246	1	KLK3_MOUSE	GLANDULAR KALLIKREIN 3	2.59e-05
19	78	42.6	259	1	KLK3_MOUSE	GLANDULAR KALLIKREIN 3	2.59e-05
20	78	42.6	261	1	KLK3_MOUSE	GLANDULAR KALLIKREIN 3	2.59e-05
21	77	42.1	259	1	KLK3_MOUSE	GLANDULAR KALLIKREIN 3	2.59e-05
22	76	41.5	256	1	KLK4_MOUSE	GLANDULAR KALLIKREIN 4	1.67e-03
23	75	41.0	259	1	KLK4_MOUSE	GLANDULAR KALLIKREIN 4	1.67e-03

24	72	39.3	232	1	KLK_PIG	GLANDULAR KALLIKREIN (	1.23e-02
25	71	38.3	241	1	TRYPS_GADMO	TRYPSINOGEN X PRECURSO	1.23e-02
26	72	39.3	241	1	TRYPS_GADMO	TRYPSINOGEN X PRECURSO	1.23e-02
27	69	37.7	241	1	TRYPS_GADMO	TRYPSINOGEN X PRECURSO	1.23e-02
28	69	37.7	241	1	TRYPS_GADMO	TRYPSINOGEN X PRECURSO	1.23e-02
29	68	37.2	239	1	KLK2_CAVPO	GLANDULAR KALLIKREIN,	5.29e-02
30	68	37.2	239	1	KLK2_CAVPO	GLANDULAR KALLIKREIN,	5.29e-02
31	68	37.2	239	1	KLK2_CAVPO	GLANDULAR KALLIKREIN,	5.29e-02
32	67	36.6	250	1	PSH_HUMAN	PROSTASIN PRECURSOR (E	8.51e-02
33	67	36.6	250	1	PSH_HUMAN	PROSTASIN PRECURSOR (E	8.51e-02
34	67	36.6	251	1	SCCE_HUMAN	TRYPSIN PRECURSOR (EC	1.36e-01
35	67	36.6	251	1	SCCE_HUMAN	TRYPSIN PRECURSOR (EC	1.36e-01
36	66	36.1	238	1	KLK_MOUSE	GLANDULAR KALLIKREIN K	1.36e-01
37	66	36.1	238	1	KLK_MOUSE	GLANDULAR KALLIKREIN K	1.36e-01
38	65	35.5	1634	1	YER6_SCHPO	HYPOPHOSPHATE RECEPTOR PRE	2.18e-01
39	64	35.0	352	1	DPOI_METJA	DNA POLYMERASE (EC 2.7	3.46e-01
40	64	35.0	352	1	DPOI_METJA	DNA POLYMERASE (EC 2.7	3.46e-01
41	63	34.4	135	1	GPDA_TREPA	GLYCEROL-3-PHOSPHATE PYROP	5.47e-01
42	63	34.4	135	1	GPDA_TREPA	GLYCEROL-3-PHOSPHATE PYROP	5.47e-01
43	63	34.4	256	1	HYPE_HYPLI	HYPOTHETICAL PROTEIN D	8.60e-01
44	63	34.4	250	1	TRY2_DROME	TRYPSIN ZETA PRECURSOR	8.60e-01
45	63	34.4	818	1	PLD2_ORYSA	PHOSPHOLIPASE D 2 PREC	8.60e-01
			1309	1	STI6_SCHPO	STE16 PROTEIN.	8.60e-01

## ALIGNMENTS

RESULT 1  
 ID PROS\_HUMAN STANDARD: PRT; 261 AA.  
 AC P07288;  
 DT 01-APR-1988 (Rel. 07, Created)  
 DT 01-JUL-1989 (Rel. 11, last sequence update)  
 DT 15-JUL-1999 (Rel. 38, last annotation update)  
 DE PROSTATE SPECIFIC ANTIGEN PRECURSOR (EC 3.4.21.77) (PSA) (GAMMA-  
 DE SEMINOPROTEIN) (SEMINOGLASE) (SEMININ) (P-30 ANTIGEN).  
 GN APS.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 CC Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-PROSTATE;  
 RX MEDLINE: 89183632.  
 RA Digby M.R., Zhang X.Y., Richard R.L.;  
 RT "Human prostate specific antigen (PSA) gene: structure and linkage to  
 RL the kallikrein-like gene, hKGL1.";  
 RL Nucleic Acids Res. 17:2137-2137(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 89282407.  
 RA Klobbeck H.-G., Compiatto G., Schulz P., Arbusow V., Fittler F.;  
 RT "Genomic sequence of human prostate specific antigen (PSA).";  
 RL Nucleic Acids Res. 17:3981-3981(1989).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-LEUKOCYTE;  
 RX MEDLINE: 89302090.  
 RA Lundwall A.;  
 RT "Characterization of the gene for prostate-specific antigen, a human  
 RL glandular kallikrein.";  
 RL Biochem. Biophys. Res. Commun. 161:1151-1159(1989).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-PROSTATE;  
 RX MEDLINE: 87190978.  
 RA Lundwall A., Lilja H.;  
 RT "Molecular cloning of human prostate specific antigen cDNA.";  
 RL FEBS Lett. 214:317-322(1987).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-PROSTATE;  
 RX MEDLINE: 89246551.  
 RA Henttu P., Viikio P.;  
 RT "cDNA coding for the entire human prostate specific antigen shows

high homologues to the human tissue kallikrein genes.";  
 Biochem. Biophys. Res. Commun. 160:903-910(1989).  
 [6]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-PROSTATE;  
 RX MEDLINE: 89165891.  
 RA Riegman P.H.J., Vlietstra R.J., van der Korput J.A.G.M., Romijn J.C.,  
 Trapman J., Vlietstra R.J., van der Korput J.A.G.M., Romijn J.C.,  
 "Characterization of the prostate-specific antigen gene: a novel  
 human kallikrein-like gene.";  
 RL Biochem. Biophys. Res. Commun. 159:95-102(1989).  
 [7]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-PROSTATE;  
 RA Wonne M.M., Moreno J.M., Mele C.M., Mulholland G.M., Gomella L.G.;  
 Submitted (Dec-1994) to the EMBL/Genbank/DBJ databases.  
 [8]  
 RP SEQUENCE OF 17-261 FROM N.A.  
 RX MEDLINE: 88289366.  
 RA Schulz P., Stucka R., Feldmann H., Combratio G., Klobeck H.-G.,  
 Fittler F.,  
 "Sequence of a cDNA clone encompassing the complete mature human  
 prostate specific antigen (PSA) and an unsplited leader sequence.";  
 RL Nucleic Acids Res. 16:6226-6226(1988).  
 [9]  
 RP SEQUENCE OF 25-261.  
 RX MEDLINE: 86205857.  
 RA Watt K.W.K., Lee P.J., W'Jinkulu T., Chan W.P., Loo R.;  
 "Human prostate-specific proteases: structural and functional  
 similarity with serine proteases.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 83:3166-3170(1986).  
 [10]  
 RP SEQUENCE OF 25-261.  
 RX MEDLINE: 88082806.  
 RA Schaller J., Akiyama K., Tsuda R., Hara M., Marti T., Rickli E.E.;  
 "Isolation, characterization and amino-acid sequence of gamma-  
 semiothionin, a glycoprotein from human seminal plasma.";  
 RL Eur. J. Biochem. 170:111-120(1987).  
 [11]  
 RP 3D-STRUCTURE MODELING.  
 RX MEDLINE: 95218633.  
 RA Villoutreix B.O., Gettoff E.D., Griffin J.H.;  
 "A structural model for the prostate disease marker, human prostate-  
 specific antigen.";  
 RL Protein Sci. 3:2033-2044(1994).  
 [12]  
 RP 3D-STRUCTURE MODELING.  
 RX MEDLINE: 98427950.  
 RA Coombs G.S., Bergstrom R.C., Pellegrer J.L., Baker S.I., Navre M.,  
 Smith M.M., Tainer J.A., Madison E., Corey D.R.;  
 "Substrate specificity of prostate-specific antigen (PSA)."  
 RL Chem. Biol. 5:475-488(1998).  
 CC -1- FUNCTION: PRESUMABLY HYDROLYZE THE HIGH MOLECULAR MASS SEMINAL  
 VESICLE PROTEIN THUS LEADING TO THE LIQUEFACTION OF THE SEMINAL  
 COAGULUM.  
 CC -1- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE: TYR-1-XAA.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. ALSO KNOWN AS THE  
 TRYPsin FAMILY. GLANDULAR KALLIKREIN SUBFAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: X05332; CAA28947.1; -  
 DR EMBL: X07730; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL: M27274; AAA60192.1; -  
 DR EMBL: M26663; AAA58802.1; -  
 DR EMBL: U17040; AAA56764.1; -  
 DR EMBL: M24543; AAA60193.1; ALT\_SEQ.  
 DR PIR: S03604; S03604.  
 DR PIR: A26757; A26757.  
 DR PIR: A32297; A32297.  
 DR PIR: A32423; A32423.  
 DR PDB: 1PSA; 26-JAN-95.  
 DR PDB: 2PSA; 16-MAR-99.  
 DR MIM: 176820; -  
 DR PRINTS: PR00722; CHYMOTRYPSIN.  
 DR PROSITE: PS00134; TRYPsin\_HIS; 1.  
 DR PROSITE: PS00135; TRYPsin\_SER; 1.  
 DR PFAM: PF00089; trypsin; 1.  
 KW Hydrolase; Serine protease; Glycoprotein; Antigen; Zymogen; Signal;  
 3D-structure.  
 FT SIGNAL 1 18  
 FT PROPEP 16 24  
 FT CHAIN 25 261  
 FT ACT\_SITE 65 65  
 FT ACT\_SITE 120 120  
 FT ACT\_SITE 213 213  
 FT CARBOHYD 69 69  
 FT DISULFID 31 173  
 FT DISULFID 50 66  
 FT DISULFID 152 219  
 FT DISULFID 184 198  
 FT DISULFID 209 234  
 FT CONFLICT 64 64  
 FT CONFLICT 73 73  
 FT CONFLICT 86 86  
 FT CONFLICT 94 94  
 FT CONFLICT 136 136  
 FT CONFLICT 165 168  
 FT CONFLICT 175 175  
 FT CONFLICT 175 175  
 FT CONFLICT 184 184  
 FT CONFLICT 260 260  
 SQ SEQUENCE 261 AA; AE9E732AF872141A CRC64;  
 Query Match 100.0%; Score 183; DB 1; Length 261;  
 Best Local Similarity 100.0%; Pred. No. 4.77e-32;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 DB 163 EEFLEPKKLCQVDLHVISNDVCAQV 187  
 QY 1 EEFLEPKKLCQVDLHVISNDVCAQV 25  
 ID PROS\_MACMU STANDARD; PRT; 261 AA.  
 AC P33619;  
 DT 01-FEB-1994 (rel. 28, Created)  
 DT 01-FEB-1994 (rel. 28, Last sequence update)  
 DT 01-OCT-1996 (rel. 34, Last annotation update)  
 DE PROSTATE SPECIFIC ANTIGEN PRECURSOR (EC 3.4.21.35) (PSA) (GAMMA-  
 SEMINOPROTEIN).  
 DE AFS.  
 GN Macaca mulatta (Rhesus macaque).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 CC Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae;  
 CC Macaca.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 93363642.  
 RA Gauthier E.R., Chapdelaine P., Tremblay R.R., Dube J.Y.;  
 "Characterization of rhesus monkey prostate specific antigen cDNA.";  
 RL Biochim. Biophys. Acta 1174:207-210(1993).  
 CC -1- FUNCTION: PSA PRESUMABLY HYDROLYZE THE HIGH MOLECULAR MASS  
 SEMINAL VESICLE PROTEIN THUS LEADING TO THE LIQUEFACTION OF

1  
 CC THE SEMINAL COAGULUM. PSA IS A GLANDULAR KALLIKREIN.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
 CC TRYPSIN FAMILY. GLANDULAR KALLIKREIN SUBFAMILY.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X73560; CAA51957.1; -  
 CC PIR: S34239; S34239.  
 CC HSSP: P07288; 1PFA.  
 CC PRINTS: PR00722; CHYMOTRYPSIN.  
 CC PROSITE: PS00134; TRYPSIN\_HIS: 1.  
 CC PROSITE: PS00135; TRYPSIN\_SER: 1.  
 CC PFM: PFO0089; trypsin; 1.  
 CC Hydrolyase; Serine protease; Glycoprotein; Antigen; Zymogen; Signal.  
 CC SIGNAL  
 CC PROPEP 19 24  
 CC ACT-1 25 261  
 CC ACT-SITE 65 65  
 CC ACT-SITE 120 120  
 CC ACT-SITE 213 213  
 CC DISULFID 31 173  
 CC DISULFID 50 66  
 CC DISULFID 152 219  
 CC DISULFID 184 198  
 CC DISULFID 209 224  
 CC CARBOHYD 102 102  
 CC SEQUENCE 261 AA; 28816 MW; 8525B14B159675C CRC64;  
 SO  
 Query Match  
 Best Local Similarity 91.3%; Score 167; DB 1; Length 261;  
 Matches 23; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 Db 163 EEFLLPKKLCQVLDHVISNDVCAQV 187  
 QY 1 EEFLLPKKLCQVLDHVISNDVCAQV 25  
 RESULT 3  
 ID KLR2\_HUMAN STANDARD: PRT; 261 AA.  
 AC P20151;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE GLANDULAR KALLIKREIN 2 PRECURSOR (EC 3.4.21.35) (TISSUE KALLIKREIN)  
 DE (PROSITE) (H0K-1).  
 GN KLR2.  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 CC Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 88054467.  
 RA Schmedlich L.J., Bennetts B.H., Morris B.J.;  
 RA "Primary structure of a human glandular kallikrein gene."  
 RA DNA 6:429-437(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE-PROSITE;  
 RA MEDLINE: 92324494.  
 RA Riegman P.H., Vileitner R.J., der Koper H.A., Romijn J.C.,  
 RA "Identification and androgen-regulated expression of two major human  
 RA glandular kallikrein-1 (hkg-1) mRNA species."  
 RA Mol. Cell. Endocrinol. 76:181-190(1991).  
 CC -1- FUNCTION: GLANDULAR KALLIKREINS CLEAVE MET-LYS AND ARG-SER BONDS  
 CC IN KININOGEN TO RELEASE LYS-BRADYKININ.

CC -1- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE OF ARG-1-XAA BONDS IN  
 CC SMALL MOLECULE SUBSTRATES. HIGHLY SELECTIVE ACTION TO RELEASE  
 CC KALLIDIN (LYSL-BRADYKININ) FROM KININOGEN INVOLVES HYDROLYSIS OF  
 CC MET-1-XAA OR LEU-1-XAA.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
 CC TRYPSIN FAMILY. GLANDULAR KALLIKREIN SUBFAMILY.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: M18156; NOT\_ANNOTATED\_CDS.  
 CC PIR: S39329; AAD13816.1; -  
 CC HSSP: A29586; A29586.  
 CC PRINTS: PR00722; CHYMOTRYPSIN.  
 CC PROSITE: PS00134; TRYPSIN\_HIS: 1.  
 CC PROSITE: PS00135; TRYPSIN\_SER: 1.  
 CC PFM: PFO0089; trypsin; 1.  
 CC Hydrolyase; Serine protease; Kininogenase; Glycoprotein;  
 CC Multigene family; Zymogen; Signal.  
 CC SIGNAL  
 CC PROPEP 19 24  
 CC ACT-1 25 261  
 CC ACT-SITE 65 65  
 CC ACT-SITE 120 120  
 CC ACT-SITE 213 213  
 CC DISULFID 31 172  
 CC DISULFID 30 66  
 CC DISULFID 152 219  
 CC DISULFID 184 198  
 CC DISULFID 209 234  
 CC CARBOHYD 102 102  
 CC SEQUENCE 261 AA; 28671 MW; 9CF7FA1102EF42D CRC64;  
 SO  
 Query Match  
 Best Local Similarity 71.0%; Score 130; DB 1; Length 261;  
 Matches 16; Conservative 5; Mismatches 3; Indels 0; Gaps 0;  
 Db 163 EEFLLPKKLCQVLDHVISNDVCAQV 186  
 QY 1 EEFLLPKKLCQVLDHVISNDVCAQV 24  
 RESULT 4  
 ID KLR1\_MACFA STANDARD: PRT; 257 AA.  
 AC 007276;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE GLANDULAR KALLIKREIN 1 PRECURSOR (EC 3.4.21.35) (TISSUE KALLIKREIN)  
 DE (KIDNEY/PANCREAS/SALIVARY GLAND KALLIKREIN).  
 GN KLR1.  
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 CC Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae;  
 CC Macaca.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 93305727.  
 RA Lin F.K., Lin C.H., Chou C., Chen K., Lu H.S., Bachelier B.,  
 RA Herrera C., Jones T., Chao J., Chao L.;  
 RA "Molecular cloning and sequence analysis of the monkey and human  
 RA tissue kallikrein genes."  
 RA Blochim. Biophys. Acta 1173:325-328(1993).  
 CC -1- FUNCTION: GLANDULAR KALLIKREINS CLEAVE MET-LYS AND ARG-SER BONDS  
 CC IN KININOGEN TO RELEASE LYS-BRADYKININ.

CC	1	-	CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE OF ARG-1-XAA BONDS IN SMALL MOLECULE SUBSTRATES. HIGHLY SELECTIVE ACTION TO RELEASE KALLIDIN (LYSL-BRADYKININ) FROM KININGEN INVOLVES HYDROLYSIS OF MET-1-XAA OR LEU-1-XAA.
CC	2	-	SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPsin FAMILY. GLANDULAR KALLIKREIN SUBFAMILY.
CC	3	-	THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/sib-sib.ch">http://www.isb-sib.ch/announce/sib-sib.ch</a> ).
CC	4	-	send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> .
CC	5	-	EMBL: L10039; AAA36853.1.
DR	6	PIR: S33772; S33772.	
DR	7	HSSP; P00752; LHIA.	
DR	8	PRINTS; PR00722; CHYMOTRYPSIN.	
DR	9	PROSITE; PS00134; TRYPSIN_HIS; 1.	
DR	10	PROSITE; PS00135; TRYPSIN_SER; 1.	
DR	11	PFAM; PF00089; trypsin; 1.	
KW	12	Hydrolase; Serine protease; Kallinogenase; Glycoprotein; Pancreas; Zymogen; Signal.	
FT	13	SIGNAL	1 18 BY SIMILARITY.
FT	14	PROPE	19 24 ACTIVATION PEPTIDE (PROBABLE).
FT	15	CHAIN	25 257 GLANDULAR KALLIKREIN 1.
FT	16	ACT_SITE	62 62 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	17	ACT_SITE	116 116 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	18	ACT_SITE	209 209 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	19	CARBOHYD	90 90 POTENTIAL.
FT	20	CARBOHYD	99 99 POTENTIAL.
FT	21	CARBOHYD	101 101 POTENTIAL.
FT	22	CARBOHYD	105 105 POTENTIAL.
FT	23	CARBOHYD	160 160 POTENTIAL.
FT	24	CARBOHYD	162 162 POTENTIAL.
FT	25	DISULFID	31 169 BY SIMILARITY.
FT	26	DISULFID	47 63 BY SIMILARITY.
FT	27	DISULFID	148 215 BY SIMILARITY.
FT	28	DISULFID	180 194 BY SIMILARITY.
FT	29	DISULFID	205 230 BY SIMILARITY.
SO	30	SEQUENCE	257 AA; 28237 MW; 32774D4C069316A7 CRC64;
Query Match	31	52.5%;	Score 96; DB 1; Length 257;
Best Local Similarity	32	56.5%;	Pred. No. 3.50e-08;
Matches	33	13; Conservative	4; Mismatches 6; Indels 0; Gaps 0;
Db	34	159 ENFSFDDLCVDELLIPNDECA 181	
OY	35	1 EEFLLPKRLGCVDLHVSNDVCA 23	
RESULT	36	5	
ID	37	KLK1_PAPHA	STANDARD; PRT; 258 AA.
AC	38	Q28773;	
DT	39	15-JUL-1998 (Rel. 36, Created)	
DT	40	15-JUL-1998 (Rel. 36, Last sequence update)	
DT	41	15-JUL-1998 (Rel. 36, Last annotation update)	
DE	42	GLANDULAR KALLIKREIN 1 PRECURSOR (EC 3.4.21.35) (TISSUE KALLIKREIN) (KIDNEY/PANCREAS/SALIVARY GLAND KALLIKREIN).	
GN	43	KLK1.	
OS	44	Eupo hamadryas (Hamadryas baboon).	
OC	45	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae; Papio.	
RN	46	[1]	
RP	47	SEQUENCE FROM N.A.	
RC	48	TISSUE-PANCREAS;	
SA	49	Perelgina L.M., Kammerer C.M., Henkel R.D.;	
RL	50	Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.	
CC	51	- FUNCTION: GLANDULAR KALLIKREINS CLEAVE MET-LYS AND ARG-SER BONDS IN KININOGEN TO RELEASE LYS-BRADYKININ.	
CC	52	- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE OF ARG-1-XAA BONDS IN	

```

CC SMALL MOLECULE SUBSTRATES. HIGHLY SELECTIVE ACTION TO RELEASE
CC KALLIDIN (LYSL-BRADYKININ) FROM KININOGEN INVOLVES HYDROLYSIS OF
CC MET-1-XAA OR LEU-1-XAA.
CC
CC -1- SIMILARITY. BELONGS TO PEPTIDASE FAMILY S1: ALSO KNOWN AS THE
CC TRYPSIN FAMILY. GLANDULAR KALLIKREIN SUBFAMILY.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its use
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC -----
CC CC EMBL: LA4121; AAA73523.1; -.
CC DR HSSP: P00752; 2KAL.
CC DR PRINTS: PRO0722; CHYMOTRYPSIN.
CC DR PROSITE: PS00134; TRYPSIN_HIS; 1.
CC DR PROSITE: PS00135; TRYPSIN_SER; 1.
CC DR PFAM: PF00089; trypsin; 1.
CC KM Hydroxase; Serine protease; Kininogenase; Glycoprotein; Pancreas;
CC FT Zymogen; Signal.
CC FT SIGNAL 1 18 BY SIMILARITY.
CC FT PROPEP 19 24 ACTIVATION PEPTIDE (PROBABLE).
CC FT CHAIN 25 258 GLANDULAR KALLIKREIN 1.
CC FT ACT_SITE 62 62 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC FT ACT_SITE 117 117 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC FT ACT_SITE 210 210 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC FT DISULFD 31 170 BY SIMILARITY.
CC FT DISULFD 47 63 BY SIMILARITY.
CC FT DISULFD 149 216 BY SIMILARITY.
CC FT DISULFD 181 195 BY SIMILARITY.
CC FT DISULFD 206 231 BY SIMILARITY.
CC FT CARBOHYD 90 90 POTENTIAL.
CC FT CARBOHYD 99 99 POTENTIAL.
CC FT CARBOHYD 101 101 POTENTIAL.
CC FT CARBOHYD 105 105 POTENTIAL.
CC FT CARBOHYD 161 161 POTENTIAL.
CC SQ SEQUENCE 258 AA; 28339 MW; 8254BEBDC84F36E4 CRC64;

Qy Query Match 51.9%; Score 95; DB 1; Length 258;
Best Local Similarity 56.5%; Pred. No. 6,15e-08;
Matches 13; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Db 160 ENFSYPDLCQCVDKILIPNDKCA 182
1 EEFLPKRLQCVDLHVISNDVCA 23

RESULT 6
ID KIKI_HUMAN STANDARD; PRT; 262 AA.
AC P06870;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE GLANDULAR KALLIKREIN 1 PRECURSOR (EC 3.4.21.35) (TISSUE KALLIKREIN)
DE (KIDNEY/PANCREAS/SALIVARY GLAND KALLIKREIN).
GN KIKI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Homiidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-PANCREAS;
RC MEDLINE: 86131629.
RA Fukushima D., Kitamura N., Nakanishi S.;
RT "Nucleotide sequence of cloned cDNA for human pancreatic kallikrein.";
RL Biochemistry 24:8037-8043(1985).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-KIDNEY;
RC MEDLINE: 88269498.
RA Evans B.A., Yun Z.X., Close J.A., Tregear G.W., Kitamura N.,

```

RA Nakanishi S., Callen D.F., Baker E., Hyland V.J., Sutherland G.R.,  
RA Richards R.I.:  
RT "Structure and chromosomal localization of the human renal kallikrein  
RT gene".  
RL Biochemistry 27:3124-3129(1988).  
RN [3]  
RN  
RP SEQUENCE FROM N.A.  
RC TISSUE=Salivary Gland;  
RX MEDLINE: 90073574.  
RA Angermann A., Bergmann C., Appelhans H.:  
RT "Cloning and expression of human salivary-gland kallikrein in  
RT Escherichia coli.".   
RL Biochem. J. 262:787-793(1989).  
RN [4]  
RN  
RP SEQUENCE OF 17-262 FROM N.A.  
RC TISSUE=KIDNEY;  
RX MEDLINE: 86135264.  
RA Baker A.R., Shine J.:  
RT "Human kidney kallikrein: cDNA cloning and sequence analysis.";  
RL DNA 4:445-450(1985).  
RN [5]  
RN  
RP SEQUENCE OF 25-262.  
RC TISSUE=URINE;  
RX MEDLINE: 89326688.  
RA Lu H.S., Lin F.-K., Chao L., Chao J.:  
RT "Human urinary kallikrein. Complete amino acid sequence and sites of  
RT glycosylation".  
RL Int. J. Pept. Protein Res. 33:237-249(1989).  
RN [6]  
RN  
RP SEQUENCE OF 25-262, AND CARBOHYDRATE-BINDING SITES.  
RC TISSUE=URINE;  
RX MEDLINE: 88203586.  
RA Kellermann J., Lotsepelch F., Geiger R., Deutzmann R.:  
RT "Human urinary kallikrein -- amino acid sequence and carbohydrate  
RT attachment sites".  
RL Protein Seq. Data Anal. 1:177-182(1988).  
RN [7]  
RN  
RP SEQUENCE OF 25-55.  
RC TISSUE=URINE;  
RX MEDLINE: 80114126.  
RA Lotsepelch F., Geiger R., Henschen A., Kutzbach C.:  
RT "N-Terminal amino acid sequence of human urinary kallikrein homology  
RT with other serine proteases".  
RL Hoppe-Seyler's Z. Physiol. Chem. 360:1947-1950(1979).  
RN [8]  
RN  
RP SEQUENCE OF 28-47.  
RC TISSUE=URINE;  
RX MEDLINE: 86223893.  
RA Takahashi S., Irie A., Katayama Y., Ito K., Miyake Y.:  
RT "N-Terminal amino acid sequence of human urinary prokallikrein.";  
RL J. Biochem. 99:989-992(1986).  
RN [9]  
RN  
RP -1- FUNCTION: GLANDULAR KALLIKREINS CLEAVE MET-LYS AND ARG-SER BONDS  
CC IN KININOGEN TO RELEASE LYS-BRADYKININ.  
CC  
CC -1- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE OF ARG-1-XAA BONDS IN  
CC SMALL MOLECULE SUBSTRATES. HIGHLY SELECTIVE ACTION TO RELEASE  
CC KALLIDIN (LYSYL-BRADYKININ) FROM KININOGEN INVOLVES HYDROLYSIS OF  
CC MET-1-XAA OR LEU-1-XAA.  
CC  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
CC TRYPsin FAMILY. GLANDULAR KALLIKREIN SUBFAMILY.  
CC  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation --  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC -----  
DR EMBL: M25629; AAA36136.1; -;  
DR EMBL: M33109; AAA59435.1; -;  
DR EMBL: M33105; AAA59455.1; JOINED.  
DR EMBL: M33106; AAA59455.1; JOINED.  
DR EMBL: M33107; AAA59455.1; JOINED.  
DR

```

DR EMBL; M33108; AAA59455.1; JOINED.
DR EMBL; X13561; CAA31912.1; -.
DR PIR; A24696; KOHU.
DR HSSP; P00757; ISGF.
DR MIM; 147910; -.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PROSITE; PS00134; TRYPSIN_HIS_1.
DR PROSITE; PS00135; TRYPSIN_SER_1.
DR PFAM; PF00089; trypsin; 1.
KW Hydrolase; serine protease; kininogenase; glycoprotein;
KV Multigene family; zymogen; signal.
FT SIGNAL 1 18 PROBABLE.
FT PROPEP 19 24 ACTIVATION PEPTIDE (PROBABLE).
FT CHAIN 25 262 GLANDULAR KALLIKREIN 1.
FT ACT_SITE 65 65 CHANGE RELAY SYSTEM.
FT ACT_SITE 120 120 CHANGE RELAY SYSTEM.
FT ACT_SITE 214 214 CHANGE RELAY SYSTEM.
FT CARBOHYD 93 93 CHANGE RELAY SYSTEM.
FT CARBOHYD 102 102
FT CARBOHYD 104 104
FT CARBOHYD 108 108
FT CARBOHYD 165 165 PARTIAL (35% OF THE CHAINS).
FT CARBOHYD 167 167 BY SIMILARITY.
FT DISULFID 31 174 BY SIMILARITY.
FT DISULFID 50 66 BY SIMILARITY.
FT DISULFID 153 220 BY SIMILARITY.
FT DISULFID 185 199 BY SIMILARITY.
FT DISULFID 210 235 BY SIMILARITY.
FT VARIANT 145 145 Q -> E.
FT FT /FTid=VAR_006625.
FT FT E -> K.
FT VARIANT 186 186 /FTid=VAR_006626.
SQ SEQUENCE 262 AA; 28889 MM; 772AD14A3A96C0 CRC64;
Query Match 49.7%; Score 91; DB 1; Length 262;
Best Local Similarity 54.5%; Pred. No. 5,75e-07;
Matches 12; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
Db 164 ENFSPPDLCVCLKTLPNDEC 165
| : | |||||:::|||
Qy 1 EEFLPKKLCCVDLHVINDVC 22

RESULT 7 STANDARD: PRT: 263 AA.
ID ID KLRK.PRAWA STANDARD: PRT: 263 AA.
AC AC P32824;
DT DT 01-OCT-1993 (Rel. 27, Created)
DT DT 01-OCT-1993 (Rel. 27, Last sequence update)
DE DE 01-NOV-1995 (Rel. 32, Last annotation update)
DE DE GLANDULAR KALLIKREIN, RENAL PRECURSOR (EC 3.4.21.35) (TISSUE
DE KALLIKREIN).
OS Osomys natelensis (African soft-furred rat) (Mastomys natalensis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mastomys.
[1]
RN RN SEQUENCE FROM N.A.
RP RP TISSUE-SALIVARY GLAND;
RC RC MEDLINE: 94226702.
RX RX Fahnstock M.;
RA RA "Characterization of kallikrein cDNAs from the African rodent
RA Mastomys.";
RL RL DNA Cell Biol. 13:293-300(1994).
CC CC -I- FUNCTION: GLANDULAR KALLIKEINS CLEAVE MET-LYS AND ARG-SER BONDS
CC CC IN KININOGEN TO RELEASE LYS-BRODKININ.
CC CC -I- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE OF ARC-I-XAA BONDS IN
CC CC SMALL MOLECULE SUBSTRATES. HIGHLY SELECTIVE ACTION TO RELEASE
CC CC KALLIDIN (LISYL-BRODKININ) FROM KININOGEN INVOLVES HYDROLYTICS OF
CC CC MET-I-XAA OR LEU-I-XAA.
CC CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC CC TRYPsin FAMILY. GLANDULAR KALLIKREIN SUBFAMILY.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the European Bioinformatics

```



CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: X17352; CAA35232.1; -  
 DR PIR: S15686; S15686.  
 DR HSSP: P00757; 1SGF.  
 DR PRINTS: PR00722; CHYMOTRYPSIN.  
 DR PROSITE: PS00134; TRYPSIN\_HIS; 1.  
 DR PFM: PFO0089; TRYPSIN\_SER; 1.  
 DR Hydrolase; Serine protease; Kinninogenase; Glycoprotein;  
 KW Multigene family; zymogen; signal.  
 FT SIGNAL 1  
 FT PROPEP 19 24 ACTIVATION PEPTIDE (PROBABLE).  
 FT CHAIN 25 263 GLANDULAR KALLIKREIN, RENAL.  
 FT ACT\_SITE 65 65 CHARGE RELAY SYSTEM.  
 FT ACT\_SITE 121 121 CHARGE RELAY SYSTEM.  
 FT ACT\_SITE 215 215 CHARGE RELAY SYSTEM.  
 FT DISULFID 31 175 BY SIMILARITY.  
 FT DISULFID 50 66 BY SIMILARITY.  
 FT DISULFID 153 221 BY SIMILARITY.  
 FT DISULFID 186 200 BY SIMILARITY.  
 FT DISULFID 211 236 BY SIMILARITY.  
 FT CARBOHYD 102 102 PROBABLE.  
 SQ SEQUENCE 263 AA; 29130 MW; A8E023B800337D5 CRC64;

Query Match 49.7%; Score 91; DB 1; Length 263;  
 Best Local Similarity 56.5%; Pred. No. 5,75e-07;  
 Matches 13; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Db 165 EEPFYSHDLCCVLELISNEVCA 187  
 ||| : |||| : ||||: ||||  
 QY 1 EEPFLPKKLCVLDHVISNDVCA 23

RESULT 8 STANDARD: PRT; 260 AA.  
 AC P09362;  
 DT 01-MAR-1989 (Rel. 10, Created)  
 DT 01-MAR-1989 (Rel. 10, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE ARGININE ESTERASE PRECURSOR (EC 3.4.21.35).  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-PROSTATE;  
 RX MEDLINE: 88211858.  
 RA Chapelaine P., Ho-Kim M.-A., Tremblay R.R., Dube J.Y.;  
 RT "Nucleotide sequence of the androgen-dependent arginine esterase mRNA  
 of canine prostate.";  
 RL FEBS Lett. 232:187-192(1988).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE: 91119675.  
 RA Chapelaine P., Gauthier E., Ho-Kim M.A., Bissonnette L.,  
 RA Tremblay R.R., Dube J.Y.;  
 RT "Characterization and expression of the prostatic arginine esterase  
 RT gene, a canine glandular kallikrein.";  
 RL DNA Cell Biol. 10:49-59(1991).  
 RN [3]  
 RP SEQUENCE OF 25-50 AND 108-145.  
 RC TISSUE-PROSTATE;  
 RX MEDLINE: 85004070.  
 RA Lazure C., Leduc R., Seidah N.G., Chretien M., Dube J.Y.,  
 RA Chapelaine P., Frenette G., Paquin R., Tremblay R.R.;  
 RT "The major androgen-dependent protease in dog prostate belongs to the  
 RT kallikrein family: confirmation by partial amino acid sequencing.";  
 RL FEBS Lett. 175:1-7(1984).

CC -1- FUNCTION: THIS SERINE PROTEASE IS FOUND IN DOG SEMINAL PLASMA,  
 CC ITS EXACT PHYSIOLOGICAL FUNCTION IS NOT KNOWN.  
 CC -1- INDUCTION: BY ANDROGENS.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
 CC TRYPSIN FAMILY. GLANDULAR KALLIKREIN SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: Y00751; CAA68720.1; -  
 DR EMBL: M63669; AAA30831.1; -  
 DR PIR: A30981; A30981.  
 DR PIR: A37938; A37938.  
 DR PIR: S00613; S00613.  
 DR HSSP: P00752; 2KAT.  
 DR PRINTS: PR00722; CHYMOTRYPSIN.  
 DR PROSITE: PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE: PS00135; TRYPSIN\_SER; 1.  
 DR PFM: PFO0089; TRYPSIN; 1.  
 KW Hydrolase; Serine protease; Glycoprotein; zymogen; signal.  
 FT SIGNAL 1  
 FT PROPEP 1 24 ACTIVATION PEPTIDE.  
 FT CHAIN 25 260 ARGININE ESTERASE.  
 FT ACT\_SITE 65 65 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 119 119 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 212 212 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT CARBOHYD 79 79 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT DISULFID 31 172 PROBABLE.  
 FT DISULFID 50 66 BY SIMILARITY.  
 FT DISULFID 151 218 BY SIMILARITY.  
 FT DISULFID 183 197 BY SIMILARITY.  
 FT DISULFID 208 233 BY SIMILARITY.  
 FT CONFLICT 56 N -> H (IN REF. 2).  
 SQ SEQUENCE 260 AA; 28746 MW; 48768B6FE204775A CRC64;

Query Match 48.6%; Score 89; DB 1; Length 260;  
 Best Local Similarity 52.0%; Pred. No. 1,73e-06;  
 Matches 13; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Db 162 ETIFPGSLCCVLDKLSNNOCARV 186  
 | : | ||||| : ||| : |||  
 QY 1 EEPFLPKKLCVLDHVISNDVCAQV 25

RESULT 9 STANDARD: PRT; 261 AA.  
 AC KLKA.MOUSE  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 15-FEB-2000 (Rel. 39, Last annotation update)  
 DE GLANDULAR KALLIKREIN K11 PRECURSOR (EC 3.4.21.35) (TISSUE KALLIKREIN)  
 DE (MGK-11).  
 GN K11 OR K1K-11.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BALB/C; TISSUE=LIVER;  
 RX MEDLINE: 89083511.  
 RA Drinkwater C.C., Richards R.I.;  
 RT "Sequence of mGK-11, a mouse glandular kallikrein gene.";  
 RL Nucleic Acids Res. 16:10918-10918(1988).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE: 85257431.  
 RA Evans B.A., Richards R.I.;  
 RT "Genes for the alpha and gamma subunits of mouse nerve growth factor



0S Mus musculus (Mouse).  
0C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
0C Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BALB/C; TISSUE-SALIVARY GLAND;  
RA MEDLINE; 88107594.  
RT Drinkwater C.C., Evans B.A., Richards R.I.;  
RT "Mouse glandular kallikrein genes: identification and  
RT characterization of the genes encoding the epidermal growth factor  
RT binding proteins.";  
RL Biochemistry 26:6750-6756(1987).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC MEDLINE; 88107593.  
RA Blaber M., Isaacson P.J., Bradshaw R.A.;  
RT "A complete cDNA sequence for the major epidermal growth factor  
RT binding protein in the male mouse submandibular gland.";  
RL Biochemistry 26:6742-6749(1987).  
RN [3]  
RP SEQUENCE OF 16-54 AND 70-122 FROM N.A.  
RC STRAIN-BALB/C; TISSUE-LIVER;  
RA Evans B.A., Drinkwater C.C., Richards R.I.;  
RT "Mouse glandular kallikrein genes. Structure and partial sequence  
RT analysis of the kallikrein gene locus.";  
RL J. Biol. Chem. 262:8027-8034(1987).  
CC -1- FUNCTION: GLANDULAR KALLIKREIN CLEAVE MET-LYS AND ARG-SER BONDS  
CC IN KININOMEN TO RELEASE LYS-BRADYKININ.  
CC -1- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE OF ARG-1-XAA BONDS IN  
CC SMALL MOLECULE SUBSTRATES. HIGHLY SELECTIVE ACTION TO RELEASE  
CC KALLIDIN (LYSY-BRADYKININ) FROM KININOMEN INVOLVES HYDROLISIS OF  
CC MET-1-XAA OR LEU-1-XAA.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
CC TRYPSIN FAMILY. GLANDULAR KALLIKREIN SUBFAMILY.  
CC  
CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL; M17962; AAA37541.1; -;  
DR EMBL; M17985; AAA37681.1; -;  
DR EMBL; M17983; AAA37681.1; JOINED.  
DR EMBL; M17984; AAA37681.1; JOINED.  
DR EMBL; M18588; AAA39350.1; ALT\_SEQ.  
DR EMBL; M18608; AAA39351.1; -;  
DR PIR; A29745; A29745.  
DR PIR; C29746; C29746.  
DR HSSP; P00757; 1SGF.  
DR MGD; MGI:95293; EGFBP3.  
DR PRINTS; PRO0722; CHYMOTRYPSIN.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
DR PFAM; PF00089; trypsin; 1  
KW Hydrolase; Serine protease; kininogenase; Glycoprotein;  
KW Multigene family; Zymogen; Signal.  
KW SIGNAL  
FT 1  
FT PROPEP 19 24  
FT CHAIN 25 261  
FT ACT\_SITE 65 65  
FT ACT\_SITE 120 120  
FT ACT\_SITE 213 213  
FT DISULFID 31 173  
FT DISULFID 30 66  
FT DISULFID 152 219  
FT DISULFID 184 198  
FT DISULFID 209 234  
FT CARBOHYD 102 102  
FT SEQUENCE 261 AA: 28900 MW: 79600.00  
FT 261 AA: 28900 MW: 79600.00



RT "Organization and expression of the rat kallikrein gene family."  
 RL J. Biol. Chem. 264:7653-7662(1989).  
 RN (3)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 89375248.  
 RA Shai S.Y., Woodley-Miller C., Chao J., Chao L.;  
 RT "Characterization of genes encoding rat tonin and a kallikrein-like  
 RL serine protease."  
 RL Biochemistry 28:5334-5343(1989).  
 RN [4]  
 RP SEQUENCE OF 25-259.  
 RX MEDLINE: 87271223.  
 RA Lazure C., Leduc R., Seldah N.G., Thibault G., Genest J., Chretien M.;  
 RT "The complete amino acid sequence of rat submaxillary gland tonin  
 RT does contain the aspartic acid at the active site: confirmation by  
 RT protein sequence analysis."  
 RL Biochem. Cell Biol. 65:321-337(1987).  
 RN [5]  
 RP SEQUENCE OF 25-103 AND 120-259.  
 RX MEDLINE: 84117504.  
 RA Lazure C., Leduc R., Seldah N.G., Thibault G., Genest J., Chretien M.;  
 RT "Amino acid sequence of rat submaxillary tonin reveals similarities  
 RT to serine proteases."  
 RL Nature 307:555-558(1984).  
 RN [6]  
 RP SEQUENCE OF 25-34.  
 RX MEDLINE: 90147705.  
 RA Kamada M., Furuhata N., Yamaguchi T., Ikekita M., Kizuki K.,  
 RA Moriya H.;  
 RT "Observation of tissue prokallikrein activation by some serine  
 RT proteases, arginine esterases in rat submandibular gland."  
 RL Biochem. Biophys. Res. Commun. 166:231-237(1990).  
 RN [7]  
 RP SEQUENCE OF 25-50, AND CHARACTERIZATION.  
 RX MEDLINE: 92250562.  
 RA Moreau T., Brillard-Bourdet M., Bouhnik J., Gauthier F.;  
 RT "Protein products of the rat kallikrein gene family. Substrate  
 RT specificities of kallikrein rk2 (tonin) and kallikrein rk9."  
 RL J. Biol. Chem. 267:10045-10051(1992).  
 RN [8]  
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).  
 RX MEDLINE: 88011252.  
 RA Fujinaga M., James M.N.G.;  
 RT "Rat submaxillary gland serine protease, tonin. Structure solution  
 RT and refinement at 1.8-A resolution."  
 RL J. Mol. Biol. 195:373-396(1987).  
 CC -1- FUNCTION: THIS PROTEIN HAS BOTH TRYPSIN- AND CHYMOTRYPSIN-LIKE  
 CC ACTIVITIES, BEING ABLE TO RELEASE ANGIOTENSIN II FROM ANGIOTENSIN  
 CC I OR ANGIOTENSINOGEN.  
 CC -1- SUBUNIT: MONOMER.  
 CC -1- TISSUE SPECIFICITY: FOUND IN SUBMAXILLARY GLAND.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
 CC TRYPSIN FAMILY. GLANDULAR KALLIKREIN SUBFAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: M1565; AAA41466.1; -  
 DR EMBL: M23878; AAA42259.1; -  
 DR EMBL: M23877; AAA42259.1; JOINED.  
 DR EMBL: M26533; AAA42081.1; -  
 DR PIR: A00945; KORTIN.  
 DR PIR: A30971; A30971.  
 DR PIR: A34050; A34050.  
 DR PIR: C23863; C23863.  
 DR PIR: B33359; B33359.  
 DR PIR: A32340; A32340.  
 DR PDB: 1TON; 16-JAN-88.

DR PRINTS: P00722; CHYMOTRYPSIN.  
 DR PROSITE: P500134; TRYPSIN\_HIS; 1.  
 DR PROSITE: P500135; TRYPSIN\_SER; 1.  
 DR PFAM: P000089; trypsin; 1.  
 KW Hydrolase; Serine protease; Glycoprotein; Multigene family; Zymogen;  
 KW Signal; 3d-structure.  
 FT SIGNAL 1  
 FT PROPEP 19 24  
 FT CHAIN 25 259  
 FT ACT\_SITE 63 63  
 FT ACT\_SITE 118 118  
 FT ACT\_SITE 211 211  
 FT ACT\_SITE 31 171  
 FT DISULFID 48 64  
 FT DISULFID 150 217  
 FT DISULFID 182 196  
 FT DISULFID 207 232  
 FT CARBOHYD 106 106  
 FT CARBOHYD 189 189  
 FT STRAND 26 26  
 FT STRAND 29 30  
 FT TURN 33 34  
 FT TURN 37 38  
 FT STRAND 39 43  
 FT STRAND 47 54  
 FT TURN 55 56  
 FT STRAND 57 60  
 FT HELIX 62 64  
 FT STRAND 70 73  
 FT STRAND 77 77  
 FT TURN 78 79  
 FT TURN 83 84  
 FT STRAND 86 88  
 FT STRAND 90 95  
 FT TURN 97 98  
 FT TURN 116 117  
 FT STRAND 120 124  
 FT TURN 146 147  
 FT STRAND 149 154  
 FT STRAND 168 168  
 FT STRAND 170 177  
 FT HELIX 179 181  
 FT HELIX 183 186  
 FT TURN 188 189  
 FT HELIX 190 193  
 FT STRAND 194 198  
 FT TURN 200 201  
 FT STRAND 205 205  
 FT TURN 208 209  
 FT TURN 211 212  
 FT STRAND 214 217  
 FT TURN 218 219  
 FT STRAND 220 225  
 FT TURN 235 236  
 FT STRAND 239 243  
 FT HELIX 244 257  
 FT SEQUENCE 259 AA; 28248 MW; 3D6E6D0D11F926B4 CRC64;  
 SO  
 Query Match 44.88; Score 82; DB 1; Length 259;  
 Best Local Similarity 39.18; Pred. No. 7,49e-05;  
 Matches 9; Conservative 8; Mismatches 6; Indels 0; Gaps 0;  
 Db 162 EMBVSHDLOCVNIHILSNKICIE 184  
 Oy 2 EFLTPKRLQCVLDHVLISNDVCAO 24  
 RESULT 13  
 ID KIK3\_RAT STANDARD: PRT: 188 AA.  
 AC P15950;  
 DT 01-APR-1990 (rel. 14, Created)  
 DT 01-JUN-1994 (rel. 29, Last sequence update)  
 DT 01-NOV-1997 (rel. 35, Last annotation update)

DE GLANDULAR KALLIKREIN 3, SUBMANDIBULAR (EC 3.4.21.35) (TISSUE  
DE KALLIKREIN) (S1 KALLIKREIN) (RKG-3) (RSKG-50) (FRAGMENT).  
GN KLK3 OR KLK-3.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 89375248.  
RA Shai S.Y., Woodley-Miller C., Chao J., Chao L.;  
RT "Characterization of genes encoding rat tonin and a kallikrein-like  
RT serine protease.";  
RL Biochemistry 28:5334-5343(1989).  
RP SEQUENCE OF 33-188 FROM N.A.  
RX MEDLINE; 86051477.  
RA Ashley P.L., Macdonald R.J.;  
RT "Kallikrein-related mrnas of the rat submaxillary gland: nucleotide  
RT sequences of four distinct types including tonin.";  
RL Biochemistry 24:4512-4520(1985).  
CC -1- FUNCTION: GLANDULAR KALLIKREINS CLEAVE MET-LYS AND ARG-SER BONDS  
CC IN KININOGEN TO RELEASE LYS-BRADYKININ.  
CC CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE OF ARG-1-XAA BONDS IN  
CC SMALL MOLECULE SUBSTRATES. HIGHLY SELECTIVE ACTION TO RELEASE  
CC KALLIDIN (LYSYL-BRADYKININ) FROM KININOGEN INVOLVES HYDROLYSIS OF  
CC MET-1-XAA OR LEU-1-XAA.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
CC TRYPsin FAMILY. GLANDULAR KALLIKREIN SUBFAMILY.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M26534; AAA42080.1; -;  
DR EMBL; M11564; AAA41465.1; -;  
DR PIR; B23863; B23863.  
DR PIR; B32340; B32340.  
DR HSSP; P00759; 1TON.  
DR PROSITE; PS00134; TRYPsin\_HIS; PARTIAL.  
DR PROSITE; PS00135; TRYPsin\_SER; 1.  
DR PRAM; PF00089; trypsin; 1.  
KW Hydrolase; Serine protease; Kininogenase; Glycoprotein;  
KW Multigene family.  
FT NON\_TER 1  
FT ACT\_SITE 47 1 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 140 140 CHARGE RELAY SYSTEM.  
FT DISULFID 79 146 BY SIMILARITY.  
FT DISULFID 111 125 BY SIMILARITY.  
FT DISULFID 136 161 BY SIMILARITY.  
FT DISULFID 34 35 MR -> IW (IN REF. 2).  
FT CONFLICT 42 42 E -> K (IN REF. 1).  
FT CONFLICT 186 186 E -> K (IN REF. 1).  
SQ SEQUENCE 188 AA; 20986 MW; B3CECE65582B1E9 CRC64;  
Query Match 44.3%; Score 81; DB 1; Length 188;  
Best Local Similarity 52.9%; Pred. No. 1.27e-04;  
Matches 9; Conservative 5; Mismatches 3; Indels 0; Gaps 0;  
Db 95 PDDLQCVNHLHLSNEKC 111  
QY 6 PKRLQCVDLHVISNDVC 22  
RESULT 14  
ID KLK8\_RAT STANDARD; PRT; 261 AA.  
AC P36374;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)

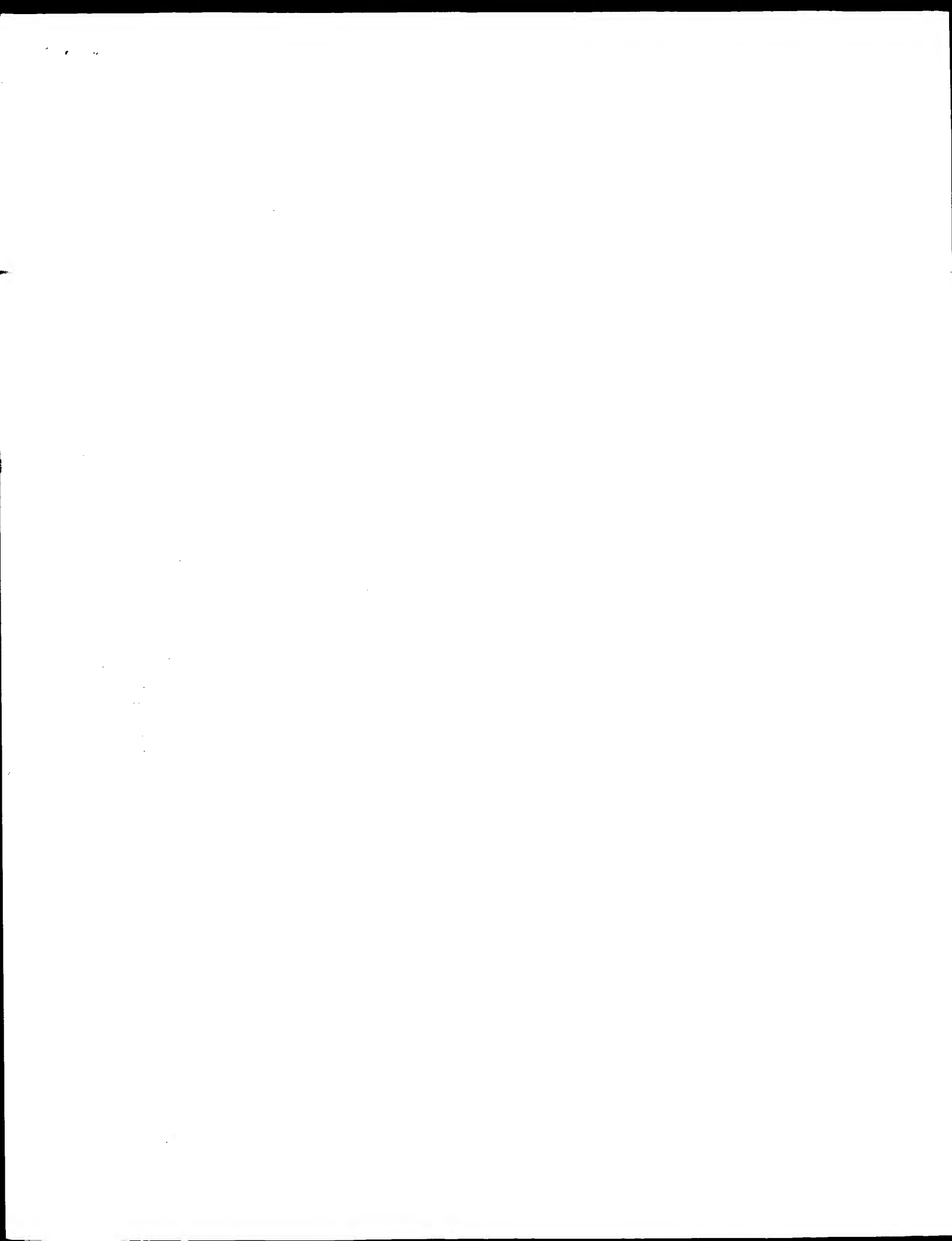
DE GLANDULAR KALLIKREIN 8, PROSTATIC PRECURSOR (EC 3.4.21.35) (TISSUE  
DE KALLIKREIN) (P1 KALLIKREIN) (RKG-8).  
GN KLK8 OR KLK-8.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 89352606.  
RA Brady J.M., Wines D.R., Macdonald R.J.;  
RT "Expression of two kallikrein gene family members in the rat  
RT prostate.";  
RL Biochemistry 28:5203-5210(1989).  
CC -1- FUNCTION: GLANDULAR KALLIKREINS CLEAVE MET-LYS AND ARG-SER BONDS  
CC IN KININOGEN TO RELEASE LYS-BRADYKININ.  
CC CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE OF ARG-1-XAA BONDS IN  
CC SMALL MOLECULE SUBSTRATES. HIGHLY SELECTIVE ACTION TO RELEASE  
CC KALLIDIN (LYSYL-BRADYKININ) FROM KININOGEN INVOLVES HYDROLYSIS OF  
CC MET-1-XAA OR LEU-1-XAA.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
CC TRYPsin FAMILY. GLANDULAR KALLIKREIN SUBFAMILY.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M27217; AAA42036.1; -;  
DR EMBL; M27215; AAA42036.1; JOINED.  
DR EMBL; M27216; AAA42036.1; JOINED.  
DR PIR; A34079; A34079.  
DR HSSP; P00759; 1TON.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PROSITE; PS00134; TRYPsin\_HIS; 1.  
DR PROSITE; PS00135; TRYPsin\_SER; 1.  
DR PRAM; PF00089; trypsin; 1.  
KW Hydrolase; Serine protease; Kininogenase; Glycoprotein;  
KW Multigene family; Zymogen; signal.  
FT SIGNAL 1 18 PROBABLE.  
FT PROPEP 19 24 ACTIVATION PEPTIDE (PROBABLE).  
FT CHAIN 25 261 GLANDULAR KALLIKREIN 8, PROSTATIC.  
FT ACT\_SITE 65 65 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 120 120 CHARGE RELAY SYSTEM.  
FT DISULFID 31 213 BY SIMILARITY.  
FT DISULFID 50 66 BY SIMILARITY.  
FT DISULFID 152 219 BY SIMILARITY.  
FT DISULFID 184 198 BY SIMILARITY.  
FT DISULFID 209 234 BY SIMILARITY.  
FT CARBOHYD 108 108 POTENTIAL.  
SQ SEQUENCE 261 AA; 29013 MW; CA8F12151B04E337 CRC64;  
Query Match 44.3%; Score 81; DB 1; Length 261;  
Best Local Similarity 52.9%; Pred. No. 1.27e-04;  
Matches 9; Conservative 5; Mismatches 3; Indels 0; Gaps 0;  
Db 168 PDDLQCVNHLHLSNEKC 184  
QY 6 PKRLQCVDLHVISNDVC 22  
RESULT 15  
ID KLK7\_RAT STANDARD; PRT; 261 AA.  
AC P36373;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE GLANDULAR KALLIKREIN 7, SUBMANDIBULAR/RENAL PRECURSOR (EC 3.4.21.35)  
DE (TISSUE KALLIKREIN) (RKG-7) (RSKG-7) (ESTERASE B) (PROTEINASE A).  
GN KLK7 OR KLK-7.

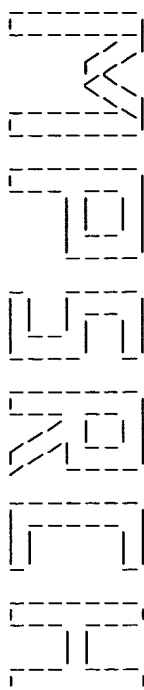
Search completed: Thu Jun 22 21:22:09 2000  
Job time : 6 secs.

OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;  
OC Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 89088074.  
RA Chen Y.-P., Chao J., Chao L.;  
RT "Molecular cloning and characterization of two rat renal kallikrein  
genes.";  
RL Biochemistry 27:7189-7196(1988).  
RN [2]  
RP SEQUENCE OF 25-75.  
RX TISSUE-SUBMAXILLARY GLAND;  
RX MEDLINE: 88198057.  
RA Kato H., Nakanishi E., Enjiyo J., Hayashi I., Oh-ishi S., Iwanaga S.;  
RT "Characterization of serine proteinases isolated from rat  
submaxillary gland: with special reference to the degradation of rat  
kininogens by these enzymes.";  
RL J. Biochem. 102:1389-1404(1987).  
CC -1- FUNCTION: GLANDULAR KALLIKREIN CLEAVE MET-LYS AND ARG-SER BONDS  
IN KININOGEN TO RELEASE LYS-BRADYKININ.  
CC -1- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE OF ARG-I-XAA BONDS IN  
SMALL MOLECULE SUBSTRATES. HIGHLY SELECTIVE ACTION TO RELEASE  
KALLIDIN (LYSYL-BRADYKININ) FROM KININOGEN INVOLVES HYDROLYSIS OF  
MET-I-XAA OR LEU-I-XAA.  
CC -1- TISSUE SPECIFICITY: KIDNEY AND SUBMANDIBULAR GLAND.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
TRYPsin FAMILY. GLANDULAR KALLIKREIN SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-slb.ch/announce/>  
or send an email to [license@isb-slb.ch](mailto:license@isb-slb.ch)).  
CC -----  
DR EMBL: M19647; AAA41461.1; -.  
DR PIR: A31136; A31136.  
DR PIR: B41429; B41429.  
DR HSSP: P00759; 1TON.  
DR PRINTS: PR00722; CHYMOTRYPSIN.  
DR PROSITE: PS00134; TRYPsin\_HIS; 1.  
DR PROSITE: PS00135; TRYPsin\_SER; 1.  
DR PIR: PF00089; trypsin; 1.  
KM Hydrolyase: Serine protease; Kininogenase; Glycoprotein;  
KW Multigene family; zymogen; signal.  
FT STGMAL 1 18  
FT PROPER 19 24 PROBABLE.  
FT CHAIN 25 261 GLANDULAR KALLIKREIN 7,  
FT ACT\_SITE 65 65 SUBMANDIBULAR/RENAL.  
FT ACT\_SITE 120 120 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 213 213 CHARGE RELAY SYSTEM.  
FT DISULFID 31 173 CHARGE RELAY SYSTEM.  
FT DISULFID 50 66 BY SIMILARITY.  
FT DISULFID 152 219 BY SIMILARITY.  
FT DISULFID 184 198 BY SIMILARITY.  
FT DISULFID 209 234 BY SIMILARITY.  
FT CARBOHYD 108 108 POTENTIAL.  
FT CONFLICT 35 35 S -> D (IN REF. 2).  
FT CONFLICT 46 46 T -> S (IN REF. 2).  
SQ SEQUENCE 261 AA: 28972 MW: 4FB06C42F25AF16 CRC64;

Query Match 44.3% Score 81; DB 1; Length 261;  
Best Local Similarity 52.9%; Pred. No. 1.2/e-04;  
Matches 9; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Db 168 PDDLQCVNHLHLSNEKC 184  
1 |||||:|:|:|  
6 PKRLQCVDLHVISNDVC 22





(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MParch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Jun 22 21:22:26 2000; Maspar time 9.42 seconds

184.051 Million cell updates/sec

Tabular output not generated.

Title: >US-09-332-866-1  
Description: (1-25) from US09332866.pep  
Perfect Score: 183  
Sequence: 1 EEFLTPKKLQCVDLHVISNDVCAOV 25

Scoring table:  
PAM 150  
Gap 15

Searched: 225878 seqs, 69334122 residues

Post-processing: Minimum Match 08  
Listing first 45 summaries

Database:

sptrembl12  
1:sp-archaea 2:sp-bacteria 3:sp-fungi 4:sp-human  
5:sp-invertebrate 6:sp-mammal 7:sp-mhc 8:sp-organelle  
9:sp-phage 10:sp-plant 11:sp-rodent 12:sp-unclassified  
13:sp-vertebrate 14:sp-virus

Statistics: Mean 28.558; Variance 35.870; scale 0.796

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	183	100.0	234	4	Q15096	PROSTATE SPECIFIC ANTI
2	130	71.0	223	4	Q15946	GLANDULAR KALLIKREIN-1
3	108	59.0	261	6	Q29474	KALLIKREIN PRECURSOR (
4	107	58.5	189	4	Q16272	PSA PROTEIN (FRAGMENT
5	92	50.3	52	6	Q46630	TISSUE KALLIKREIN (EC
6	91	49.7	195	4	Q07277	KALLIKREIN PRECURSOR
7	87	47.5	250	11	Q03955	GLANDULAR KALLIKREIN (
8	86	47.0	254	4	Q9Y5K2	KALLIKREIN 4.
9	81	44.3	233	11	Q63275	KALLIKREIN (FRAGMENT)
10	80	43.7	265	5	Q16101	SERINE PROTEASE SERP4 P
11	78	42.6	254	6	Q9XSM6	ENAMEL MATRIX SERINE P
12	78	42.6	261	11	Q61855	TISSUE KALLIKREIN.
13	76	41.5	96	11	Q61761	KALLIKREIN (FRAGMENT)
14	75	41.0	255	11	Q9Z0M1	ENAMEL MATRIX SERINE P
15	74	40.4	235	11	Q63274	KALLIKREIN (FRAGMENT)
16	74	40.4	1496	3	Q74788	CONSERVED HYPOTHETICAL
17	73	39.9	148	5	Q25069	SERINE PROTEASE (FRAGM
18	73	39.9	256	5	Q18599	DUTRY-1 TRYPSIN PRECUR
19	72	39.3	196	5	Q45044	POTATIVE SERINE PROTEI
20	72	39.3	204	5	Q17954	C14A4.7 PROTEIN.

21	72	39.3	257	5	Q97099	CHYMOTRYPSIN 2.	1.53e-02
22	72	39.3	259	5	Q97097	CHYMOTRYPSIN 1.	1.53e-02
23	71	38.8	785	3	Q75006	SUBUNIT OF THE FINAL S	2.52e-02
24	70	38.3	219	13	Q91036	TRYPSINOGEN I (FRAGMEN	4.14e-02
25	70	38.3	263	11	Q9WU18	ORIGIN RECOGNITION COM	4.14e-02
26	69	37.7	250	13	Q93265	TRYPSINOGEN 1 PRECURSO	6.76e-02
27	68	37.2	232	5	Q94508	DER F 3 MITE ALLERGEN	1.10e-01
28	68	37.2	249	13	Q92046	PREPROTRYPSIN PRECURSO	1.10e-01
29	68	37.2	396	5	Q25553	HYPOTHETICAL PROTEIN (	1.10e-01
30	67	36.6	243	5	Q01309	TRYPSINOGEN.	1.79e-01
31	67	36.6	261	11	Q88309	PRORENIN-CONVERTING EN	1.79e-01
32	67	36.6	443	5	Q19879	F8BD1.8 PROTEIN.	1.79e-01
33	66	36.1	246	11	Q88301	SERINE PROTEASE (BSP).	2.88e-01
34	66	36.1	283	4	Q95518	D1110K4.2 (NOVEL TRYP	2.88e-01
35	65	35.5	154	5	Q18448	CHYMOTRYPSIN-LIKE PROT	4.63e-01
36	65	35.5	256	13	Q93267	TRYPSINOGEN 3 PRECURSO	4.63e-01
37	65	35.5	284	5	Q01953	SERINE PROTEASE PRECUR	4.63e-01
38	65	35.5	292	5	Q18444	CHYMOTRYPSIN-LIKE PROT	4.63e-01
39	65	35.5	293	4	Q9Y337	KALLIKREIN-LIKE PROTEI	4.63e-01
40	65	35.5	295	5	Q18445	CHYMOTRYPSIN-LIKE PROT	4.63e-01
41	65	35.5	295	5	Q18450	CHYMOTRYPSIN-LIKE PROT	4.63e-01
42	65	35.5	553	14	Q9YWC6	MA55.	4.63e-01
43	65	35.5	1282	11	Q60520	PAIRED AMPHIPATHIC HEL	4.63e-01
44	65	35.5	1777	14	Q89278	NONSTRUCTURAL PROTEIN	4.63e-01
45	65	35.5	2183	14	Q64962	VIRAL RNA-DEPENDENT RN	4.63e-01

## ALIGNMENTS

RESULT	ID	1	PRELIMINARY:	PRT:	234 AA.
AC	Q15096				
DT	01-NOV-1996	(TREMBLrel. 01, Created)			
DT	01-NOV-1996	(TREMBLrel. 01, Last sequence update)			
DT	01-NOV-1999	(TREMBLrel. 12, Last annotation update)			
DE	PROSTATE SPECIFIC ANTIGEN (PA) PRECURSOR (FRAGMENT).				
GN	APS.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;				
CC	Eutheria; Primates; Catarrhini; Homnidae; Homo.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE; 88326297.				
RA	RIEGMAN P.H., KLAASSEN P., DER KOPPUT J.A., ROMIJN J.C., TRAPMAN J.;				
RT	"Molecular cloning and characterization of novel prostate antigen				
RT	cdna's."				
RL	Biochem. Biophys. Res. Commun. 155:181-188(1988).				
DR	EMBL; M21896; AAAS9996.1; -.				
DR	HSSP; P07288; IPFA.				
DR	PFAM; PF00089; trypsin; 1.				
FT	NON-TER	1	1		
FT	SIGNAL	<1	13	POTENTIAL.	
FT	CHAIN	21	234	POTENTIAL.	
SO	SEQUENCE	234 AA;	23803 MW;	1B244D8C CRC32;	
Query Match		100.0%;	Score 183;	DB 4;	Length 234;
Best Local Similarity		100.0%;	Pred. No. 3.02e-32;		
Matches	25;	Conservative	0;	Mismatches	0;
				Indels	0;
				Gaps	0;
Db	159	EEFLTPKKLQCVDLHVISNDVCAOV	183		
QY	1	EEFLTPKKLQCVDLHVISNDVCAOV	25		
RESULT	2				
ID	Q15946	PRELIMINARY:	PRT:	223 AA.	
AC	Q15946				
DT	01-NOV-1996	(TREMBLrel. 01, Created)			
DT	01-NOV-1996	(TREMBLrel. 01, Last sequence update)			
DT	01-NOV-1999	(TREMBLrel. 12, Last annotation update)			
DE	GLANDULAR KALLIKREIN-1 PROTEIN.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;				

```

OC Eutheria; Primates; Catarrhini; Homiidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 92324494.
RA RIEGMAN P.H., VILLETSTRA R.J., VAN DER KORPUT H.A., ROMIJN J.C.,
RA TRAMMAN J.,
RT "Identification and androgen-regulated expression of two major human
RT glandular kallikrein-1 (hgr-1) mRNA species."
RL Mol. Cell. Endocrinol. 76:181-190(1991).
DR EMBL: S39329; AAD13817.1; -.
DR HSSP: P07288; 1PFA.
DR PFAM: PF00089; trypsin; 1.
SQ SEQUENCE 223 AA; 24674 MW; E0321503 CRC32;

Query Match
Best Local Similarity 66.7%; Score 130; DB 4; Length 223;
Matches 16; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Db 163 EEF1PKKLCQVDLHVISNDVCA 186
QY 1 EEF1PKKLCQVDLHVISNDVCA 24

RESULT 3
ID 029474 PRELIMINARY; PRT; 261 AA.
AC 029474;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE KALLIKREIN PRECURSOR (EC 3.4.21.35) (TISSUE KALLIKREIN)
DE (GLANDULAR KALLIKREIN) (KININOGENIN).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Carnivora; Fissipedia; Canidae; Canis.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-MONGREL;
RX MEDLINE: 94250683.
RA GAUTHIER E.R., DUMAS C., CHAPPELAINE P., TREMBLAY R.R., DUBE J.Y.,
RT "Characterization of canine pancreas kallikrein cDNA."
RL Biochim. Biophys. Acta 1218:102-104(1994).
CC -1- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE OF ARG-1-XAA BONDS IN
CC SMALL MOLECULE SUBSTRATES. HIGHLY SELECTIVE ACTION TO RELEASE
CC KALLIDIN (LYSYL-BRADYKININ) FROM KININOGEN INVOLVES HYDROLYSIS OF
CC MET-1-XAA OR LEU-1-XAA. THE RAT ENZYME IS UNUSUAL IN LIBERATING
CC BRADYKININ DIRECTLY FROM AUTOLOGOUS KININOGENS BY CLEAVAGE AT TWO
CC ARG-1-XAA BONDS.
DR EMBL: X75479; CA53210.1; -.
DR HSSP: P00752; 2KAI.
DR PFAM: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
KW Signal; Hydrolyase.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 261 KALLIKREIN.
SQ SEQUENCE 261 AA; 28913 MW; 65B49AEF CRC32;

Query Match
Best Local Similarity 59.0%; Score 108; DB 6; Length 261;
Matches 13; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Db 163 DKFIYDDLCQVDLHLSNDICA 185
QY 1 EEF1PKKLCQVDLHVISNDVCA 23

RESULT 4
ID 016272 PRELIMINARY; PRT; 189 AA.
AC 016272;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE PSA PROTEIN (FRAGMENT).
OS Eukaryota; Primates; Catarrhini; Homiidae; Homo.

```

```

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Homiidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 95079406.
RA MONNE M., CROCE C.M., YU H., DIAMANDIS E.P.;
RT "Molecular characterization of prostate-specific antigen messenger RNA
RT expressed in breast tumors."
RL Cancer Res. 54:6344-6347(1994).
DR EMBL: S75755; AAD14185.1; -.
DR HSSP: P07288; 1PFA.
DR PFAM: PF00089; trypsin; 1.
FT NON_TER 1 1
SQ SEQUENCE 189 AA; 21039 MW; DBAC6482 CRC32;

Query Match
Best Local Similarity 100.0%; Score 107; DB 4; Length 189;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 176 EEF1PKKLCQVDL 189
QY 1 EEF1PKKLCQVDL 14

RESULT 5
ID 046630 PRELIMINARY; PRT; 52 AA.
AC 046630;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE TISSUE KALLIKREIN (EC 3.4.21.35) (GLANDULAR KALLIKREIN) (KININOGENIN)
DE (FRAGMENT)
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae;
RN [1]
RP SEQUENCE FROM N.A.
RX DEDIO J., WOHLEFART P., WIEMER G., SCHOLKENS B.A., MULLER-ESTERL W.;
RT Submitted (NOV-1997) to the EMBL/Genbank/DBJ databases.
RL -1- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE OF ARG-1-XAA BONDS IN
CC SMALL MOLECULE SUBSTRATES. HIGHLY SELECTIVE ACTION TO RELEASE
CC KALLIDIN (LYSYL-BRADYKININ) FROM KININOGEN INVOLVES HYDROLYSIS OF
CC MET-1-XAA OR LEU-1-XAA. THE RAT ENZYME IS UNUSUAL IN LIBERATING
CC BRADYKININ DIRECTLY FROM AUTOLOGOUS KININOGENS BY CLEAVAGE AT TWO
CC ARG-1-XAA BONDS.
DR EMBL: AJ003191; CA05970.1; -.
DR HSSP: P00752; 2KAI.
KW Hydrolyase.
FT NON_TER 1 1
FT NON_TER 52 52
SQ SEQUENCE 52 AA; 5763 MW; C9EB33B7 CRC32;

Query Match
Best Local Similarity 52.2%; Score 92; DB 6; Length 52;
Matches 12; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Db 5 DEFSYDDLCQVDLILPNEKCA 27
QY 1 EEF1PKKLCQVDLHVISNDVCA 23

RESULT 6
ID 007277 PRELIMINARY; PRT; 195 AA.
AC 007277;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE KALLIKREIN PRECURSOR (EC 3.4.21.35) (TISSUE KALLIKREIN).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Homiidae; Homo.

```



RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-LIVER:  
 RA MEDLINE: 93305727.  
 RA LIN F.K., LIN C.H., CHOU C., CHEN K., LU H.S., BACHELLER B.,  
 RA HERRERA C., JONES T., CHAO J., CHAO L.,  
 RT "Molecular cloning and sequence analysis of the monkey and human  
 RT tissue kallikrein genes."  
 RL Biochim. Biophys. Acta 1173:325-328(1993).  
 DR EMBL: L10038; AAA36135.1; -.  
 DR HSSP: P07288; 1PPA.  
 DR PFM: PF00089; trypsin: 1.  
 KW Hydrolyase: Serine protease; kininogenase; glycoprotein; zymogen;  
 KN Signal.  
 FT PROPEP 1 2 ACTIVATION PEPTIDE.  
 FT SIGNAL 1 2  
 FT CHAIN 1 2  
 FT ACT\_SITE 147 147 KALLIKREIN.  
 FT DISULFID 86 153 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT DISULFID 118 132 BY SIMILARITY.  
 FT DISULFID 143 168 BY SIMILARITY.  
 FT CARBOHYD 98 98 POTENTIAL.  
 FT CARBOHYD 100 100 POTENTIAL.  
 SQ SEQUENCE 195 AA; 21091 MW; 0DIE38EC CRC32;  
 Query Match 49.7%; Score 91; DB 4; Length 195;  
 Best Local Similarity 54.5%; Pred. No. 5.98e-07;  
 Matches 12; Conservative 4; Mismatches 6; Indels 0; Gaps 0;  
 Db 97 ENFSPPDLCVLDKILPNDK 118  
 Oy 1 EEFLLPKKLCQVLDHVSNDVC 22  
 RESULT 7  
 ID 003955 PRELIMINARY; PRT; 250 AA.  
 AC 003955;  
 DT 01-NOV-1996 (TEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)  
 DT 01-NOV-1999 (TEMBLrel. 12, Last annotation update)  
 DE GLANDULAR KALLIKREIN (FRAGMENT).  
 OS Prionys natalensis (African soft-furred rat) (Mastomys natalensis).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mastomys.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-SALIVARY GLAND;  
 RX MEDLINE: 94226702.  
 RA BELL R.A., FAHNESTOCK M.,  
 RT "Characterization of kallikrein cDNAs from the African rodent  
 RT Mastomys."  
 RL DNA Cell Biol. 13:293-300(1994).  
 CC -1- FUNCTION: GLANDULAR KALLIKREIN CLEAVE MET-LYS AND ARG-SER BONDS  
 CC IN KININOGEN TO RELEASE LYS-BRADYKININ.  
 DR EMBL: X17351; CAA35231.1; -.  
 DR HSSP: P36368; 1A05.  
 DR PFM: PF00089; trypsin: 1.  
 KW Hydrolyase; Serine protease; kininogenase; glycoprotein; Saliva;  
 KW Submandibular gland; Multigene family; Zymogen; Signal.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 250 AA; 27385 MW; BD8435E0 CRC32;  
 Query Match 47.5%; Score 87; DB 11; Length 250;  
 Best Local Similarity 47.8%; Pred. No. 5.54e-06;  
 Matches 11; Conservative 6; Mismatches 6; Indels 0; Gaps 0;  
 Db 161 EEKFPDLCVLDKILPNDK 183  
 Oy 1 EEFLLPKKLCQVLDHVSNDVC 23  
 RESULT 8  
 ID 09Y5K2 PRELIMINARY; PRT; 254 AA.

AC 09Y5K2;  
 DT 01-NOV-1999 (TEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TEMBLrel. 12, Last sequence update)  
 DT 01-NOV-1999 (TEMBLrel. 12, Last annotation update)  
 DE KALLIKREIN 4.  
 GN KLK4.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STEPHENSON S.A., VERITY K., ASHWORTH L., CLEMENTS J.A.;  
 RT "Localization of a new prostate specific antigen-related serine  
 RT protease gene, KLK4, is evidence for an expanded human kallikrein  
 RT (KLK) gene family cluster on chromosome 19q13.3-13.4";  
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF148532; AAD38019.1; -.  
 DR EMBL: AF148532; AAD38019.1; -.  
 SQ SEQUENCE 254 AA; 27022 MW; 0CFAD39C CRC32;  
 Query Match 47.0%; Score 86; DB 4; Length 254;  
 Best Local Similarity 45.0%; Pred. No. 9.60e-06;  
 Matches 9; Conservative 7; Mismatches 4; Indels 0; Gaps 0;  
 Db 162 PVLQCVNVSVEVCSTL 181  
 Oy 6 PKKLCVLDHVSNDVC 25  
 RESULT 9  
 ID 063275 PRELIMINARY; PRT; 239 AA.  
 AC 063275;  
 DT 01-NOV-1996 (TEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)  
 DT 01-NOV-1999 (TEMBLrel. 12, Last annotation update)  
 DE KALLIKREIN (FRAGMENT).  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-SPRAGUE-DAWLEY; TISSUE-SUBMANDIBULAR GLAND;  
 RA ZINTZ C.B., MA J.X., CHAO J., CHAO L.,  
 RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: L33840; AAA58782.1; -.  
 DR HSSP: P00759; 1TON.  
 DR PFM: PF00089; trypsin: 1.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 239 AA; 26382 MW; 0609E3B6 CRC32;  
 Query Match 44.3%; Score 81; DB 11; Length 239;  
 Best Local Similarity 52.9%; Pred. No. 1.44e-04;  
 Matches 9; Conservative 5; Mismatches 3; Indels 0; Gaps 0;  
 Db 146 PDLQCVNHLNNEKC 162  
 Oy 6 PKKLCVLDHVSNDVC 22  
 RESULT 10  
 ID 016101 PRELIMINARY; PRT; 265 AA.  
 AC 016101;  
 DT 01-JAN-1998 (TEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)  
 DT 01-NOV-1999 (TEMBLrel. 12, Last annotation update)  
 DE SERINE PROTEASE SERA PRECURSOR.  
 GN SER4.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-OREGON R;

RX MEDLINE: 98146263.  
 RA ARENS J.E., MAHONEY P.A.:  
 RT "Isolation and cloning of Ser4, a gene encoding a trypsin-like serine  
 RT protease in *Drosophila melanogaster*.";  
 RL Biochim. Biophys. Acta 1395:141-144 (1998).  
 DR EMBL: AF006639; AAC14351.1; -  
 DR HSSP: P08897; 1HYL.  
 DR FLYBASE: FBgn0020906; Ser4.  
 DR PFAM: PF00089; trypsin.1.  
 KM Signal; Protease.  
 FT SIGNAL 1 21  
 FT CHAIN 37 265  
 SQ SEQUENCE 265 AA; 28950 MW; 5272F05B CRC32;  
 Query Match 43.7%; Score 80; DB 5; Length 265;  
 Best Local Similarity 47.4%; Pred. No. 2,45e-04;  
 Matches 9; Conservative 6; Mismatches 4; Indels 0; Gaps 0;  
 Db 171 PDMMECVDLQIISNCSR 189  
 QY 6 PKKLQCVDLHVISNDVC 24  
 RESULT 11  
 ID 09XSN6; PRELIMINARY; PRT; 254 AA.  
 AC 09XSN6;  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DE ENAMEL MATRIX SERINE PROTEINASE 1 PRECURSOR.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA SIMMER J.P., FUKAE M., TANABE T., YAMAKOSHI Y., UCHIDA T., XUE J.,  
 RA MARGOLIS H.C., SHIMIZU M., HU C.-C., BARTLETT J.D.:  
 RT "Purification, Characterization and Cloning of Enamel Matrix Serine  
 RT Proteinase 1.";  
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U76256; AAB94638.1; -  
 KW Signal.  
 FT SIGNAL 1 24  
 FT CHAIN 31 254  
 SQ SEQUENCE 254 AA; 27235 MW; 25C6DAA2 CRC32;  
 Query Match 42.6%; Score 78; DB 6; Length 254;  
 Best Local Similarity 52.9%; Pred. No. 7.02e-04;  
 Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;  
 Db 162 PVLQCVNISVASEVC 178  
 QY 6 PKKLQCVDLHVISNDVC 22  
 RESULT 12  
 ID 061855; PRELIMINARY; PRT; 261 AA.  
 AC 061855;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DE TISSUE KALLIKREIN.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA TADA M., PETERS J., TAKAHASHI S., INOUE H., MIYAKE Y.:  
 RT "Identification of a tissue kallikrein gene, mck-6, expressed in a  
 RT mouse neuroendocrine cell line.";  
 RL Submitted (JUN-1991) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: D10464; BAA01257.1; -

DR HSSP: P36368; 1AOS.  
 DR PFAM: PF00089; trypsin.1  
 DR PRINTS: PRO0722; CHYMOTRYPSIN.  
 SQ SEQUENCE 261 AA; 28775 MW; 509D1D00 CRC32;  
 Query Match 42.6%; Score 78; DB 11; Length 261;  
 Best Local Similarity 50.0%; Pred. No. 7.02e-04;  
 Matches 9; Conservative 5; Mismatches 4; Indels 0; Gaps 0;  
 Db 168 PDELQCVNLKLPNEDCA 185  
 QY 6 PKKLQCVDLHVISNDVC 23  
 RESULT 13  
 ID 061761; PRELIMINARY; PRT; 96 AA.  
 AC 061761;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DE KALLIKREIN (FRAGMENT).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MASON A.D., EVANS B.A., COX D.R., SHINE J., RICHARDS R.I.:  
 RT "Structure of mouse kallikrein gene family suggests a role in specific  
 RT processing of biologically active peptides.";  
 RL Nature 303:300-307 (1983)  
 DR EMBL: V00829; CAA24212.1; -  
 DR HSSP: P36368; 1AOS.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 96 AA; 10327 MW; CDDC7116 CRC32;  
 Query Match 41.5%; Score 76; DB 11; Length 96;  
 Best Local Similarity 47.1%; Pred. No. 1.99e-03;  
 Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;  
 Db 3 AKDLQCVNLKLPNENC 19  
 QY 6 PKKLQCVDLHVISNDVC 22  
 RESULT 14  
 ID 0920M1; PRELIMINARY; PRT; 255 AA.  
 AC 0920M1;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DE ENAMEL MATRIX SERINE PROTEINASE 1 PRECURSOR.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-SWISS-WEBSTER;  
 RA SIMMER J.:  
 RT "Enamel Matrix Serine Proteinase 1 (EMSP1).";  
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF019979; AAC98894.1; -  
 DR HSSP: P00763; 1DPO.  
 KW Signal.  
 FT SIGNAL 1 25  
 FT CHAIN 32 255  
 SQ SEQUENCE 255 AA; 27488 MW; 6E711616 CRC32;  
 Query Match 41.0%; Score 75; DB 11; Length 255;  
 Best Local Similarity 52.9%; Pred. No. 3.33e-03;  
 Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
 Db 163 PSLQCVNLVASEENC 179

OY 6 PKKLOCVDLHVISNDVC 22

RESULT 15  
ID 063274 PRELIMINARY; PRT: 235 AA.

AC 063274;  
DT 01-NOV-1996 (Tremblrel. 01, Created)  
DT 01-NOV-1996 (Tremblrel. 01, last sequence update)  
DT 01-NOV-1999 (Tremblrel. 12, last annotation update)  
DE KALIKREIN (FRAGMENT).  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

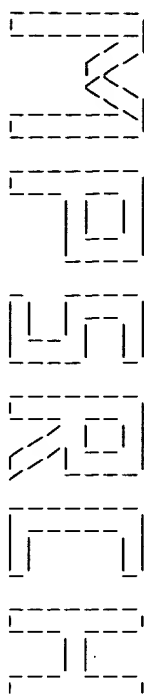
RP SEQUENCE FROM N.A.  
RC STRAIN-SPRAGUE-DAWLEY; TISSUE-KIDNEY;  
RA 2INT2 C.B., MA J.X., CHAO J., CHAO L.;  
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.  
DR EMBL; L33839; AAA58781.1; -.  
DR HSSP; P00759; ITON.  
DR PFAM; PF00089; trypsin; 1.  
FT NON-TER 1  
SQ SEQUENCE 235 AA; 26226 MW; E6DE5AD8 CRC32;

Query Match 40.4%; Score 74; DB 11; Length 235;  
Best Local Similarity 42.1%; Pred. No. 5.55e-03;  
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

DB 142 PDDLCCVNDLSNKCIE 160  
OY 6 PKKLOCVDLHVISNDVC 24

Search completed: Thu Jun 22 21:22:40 2000  
Job time : 14 secs.





(TM)

Release 3.1A John F. Collins, Bioinformatics Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsearch n.a. - n.a. Smith-Waterman search, using a protein query  
which has been backtranslated into n.a. using IUPAC symbols

Run on: Fri Jun 23 15:53:38 2000; MasPar time 4470.62 seconds

Tabular output not generated. 67.998 Million cell updates/sec

Title: >US-09-332-866-1  
Description: (1-25) from US09332866.pep  
Perfect Score: 316  
N.A. Sequence: 1 GARGARTTYTACNCCNNA.....AYGATGTNGTCNCARGTN 75  
Comp: CTCTCTAARANTGNGNTT.....TCTCTACANRCGNGTTCAN

Scoring table: TABLE bkttranslated2  
Gap 40

Mismatch STD : Dbase 0; Query 0

Searched: 4857316 seqs, 2026611650 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database:

embl-est61  
1:em\_est11 2:em\_est12 3:em\_est13 4:em\_est14 5:em\_est15  
6:em\_est16 7:em\_est17 8:em\_est18 9:em\_est19 10:em\_est20  
11:em\_est21 12:em\_est22 13:em\_est23 14:em\_est24 15:em\_est25  
16:em\_gss12 17:em\_gss2 18:em\_gss3 19:em\_gss4 20:em\_gss5  
21:em\_gss6  
genbank-est116  
22:gb\_est1 23:gb\_est10 24:gb\_est11 25:gb\_est12  
26:gb\_est13 27:gb\_est14 28:gb\_est15 29:gb\_est16  
30:gb\_est17 31:gb\_est18 32:gb\_est19 33:gb\_est20  
34:gb\_est21 35:gb\_est22 36:gb\_est23 37:gb\_est24  
38:gb\_est25 39:gb\_est26 40:gb\_est27 41:gb\_est28  
42:gb\_est29 43:gb\_est30 44:gb\_est31 45:gb\_est32  
46:gb\_est33 47:gb\_est34 48:gb\_est35 49:gb\_est36  
50:gb\_est37 51:gb\_est38 52:gb\_est39 53:gb\_est40  
54:gb\_est41 55:gb\_est42 56:gb\_est43 57:gb\_est44  
58:gb\_est45 59:gb\_est46 60:gb\_est47 61:gb\_est48  
62:gb\_est49 63:gb\_est50 64:gb\_est51 65:gb\_est52  
66:gb\_est53 67:gb\_est54 68:gb\_est55 69:gb\_est56  
70:gb\_est57 71:gb\_est58 72:gb\_est59 73:gb\_est60  
74:gb\_est61 75:gb\_est62 76:gb\_est63 77:gb\_est64  
78:gb\_est65 79:gb\_est66 80:gb\_est67 81:gb\_est68  
82:gb\_est69 83:gb\_est70 84:gb\_est71 85:gb\_est72  
86:gb\_est73 87:gb\_est74 88:gb\_est75 89:gb\_est76  
90:gb\_est77 91:gb\_est78 92:gb\_est79 93:gb\_est80  
94:gb\_est81 95:gb\_est82 96:gb\_est83 97:gb\_est84  
98:gb\_est85 99:gb\_est86 100:gb\_est87

1	312	98.7	373	24	AA228953	nc14604.r1	NCI_CGAP_Pr	1.08e-49
2	312	98.7	388	24	AA228882	nc15804.r1	NCI_CGAP_Pr	1.08e-49
3	312	98.7	420	24	AA228836	nc16106.r1	NCI_CGAP_Pr	1.08e-49
4	312	98.7	426	24	AA226359	nc17109.r1	NCI_CGAP_Pr	1.08e-49
5	312	98.7	436	30	AA630877	nc15805.s1	NCI_CGAP_Pr	1.08e-49
6	312	98.7	439	29	AA569482	nc123611.s1	NCI_CGAP_Pr	1.08e-49
7	312	98.7	453	29	AA573727	nc143609.s1	NCI_CGAP_Pr	1.08e-49
8	312	98.7	466	30	AA545277	nc159812.s1	NCI_CGAP_Pr	1.08e-49
9	312	98.7	471	30	AA654924	nc17607.s1	NCI_CGAP_Pr	1.08e-49
10	312	98.7	481	29	AA579159	nc18508.s1	NCI_CGAP_Pr	1.08e-49
11	312	98.7	489	30	AA658261	nc121003.s1	NCI_CGAP_Pr	1.08e-49
12	312	98.7	491	24	AA225115	nc121011.s1	NCI_CGAP_Pr	1.08e-49
13	312	98.7	505	29	AA574033	nc144511.s1	NCI_CGAP_Pr	1.08e-49
14	312	98.7	520	28	AA544295	nc170605.s1	NCI_CGAP_Pr	1.08e-49
15	312	98.7	542	30	AA654558	nc159811.s1	NCI_CGAP_Pr	1.08e-49
16	312	98.7	554	29	AA579039	nc134907.s1	NCI_CGAP_Pr	1.08e-49
17	312	98.7	681	47	AI732097	nc121011.x5	NCI_CGAP_Pr	1.08e-49
18	312	98.7	732	43	AI547309	PM001_AH.B02.r	Yodnorm	1.08e-49
19	312	98.7	779	43	AI547084	PM2.1.1.G05.r	mynorm	1.08e-49
20	307	97.2	272	22	T29518	EST82635	Human Prostac	2.53e-48
21	302	95.6	398	30	AA635287	nc171007.s1	NCI_CGAP_Pr	5.87e-47
22	302	95.6	457	51	AI926979	nc168011.x1	NCI_CGAP_Pr	5.87e-47
23	302	95.6	558	30	AA640352	nc121012.s1	NCI_CGAP_Pr	5.87e-47
24	292	92.4	487	24	AA229599	nc15802.s1	NCI_CGAP_Pr	3.05e-44
25	277	87.7	328	30	AA654296	nc18306.s1	NCI_CGAP_Pr	3.32e-40
26	272	86.1	406	27	AA420570	nc161006.r1	NCI_CGAP_Pr	7.18e-39
27	272	86.1	500	27	AA420845	nc160004.s1	NCI_CGAP_Pr	7.18e-39
28	257	81.3	582	29	AA594946	nc161006.s1	NCI_CGAP_Pr	6.69e-35
29	252	79.7	414	27	AA420522	nc172008.s1	NCI_CGAP_Pr	1.37e-33
30	242	76.6	310	28	AA522842	nc172008.s1	NCI_CGAP_Pr	5.45e-31
31	242	76.6	329	29	AA558309	nc135004.s1	NCI_CGAP_HS	5.45e-31
32	242	76.6	902	42	AI525832	nc135004.s1	NCI_CGAP_HS	5.45e-31
33	232	73.4	273	30	AA639308	nc196004.s1	NCI_CGAP_Pr	2.04e-28
34	232	73.4	283	28	AA535453	nc176001.s1	NCI_CGAP_Pr	2.04e-28
35	232	73.4	377	24	AA226360	nc171009.s1	NCI_CGAP_Pr	2.04e-28
36	232	73.4	804	43	AA557338	nc171009.s1	NCI_CGAP_Pr	2.04e-28
37	210	66.5	189	30	AA577663	nc182006.s1	NCI_CGAP_Pr	7.12e-23
38	191	60.4	104	30	AA657697	nc182006.s1	NCI_CGAP_Pr	3.11e-18
39	191	60.4	485	29	AA573575	nc141110.s1	NCI_CGAP_Pr	3.11e-18
40	181	57.3	413	47	AI732674	nc167005.s1	NCI_CGAP_Pr	7.42e-16
41	181	57.3	428	28	AA469129	nc167005.s1	NCI_CGAP_Pr	7.42e-16
42	181	57.3	433	29	AA551415	nc155802.s1	NCI_CGAP_Pr	7.42e-16
43	181	57.3	433	28	AA534246	nc170611.s1	NCI_CGAP_Pr	7.42e-16
44	181	57.3	537	29	AA55047	nc104003.s1	NCI_CGAP_Pr	7.42e-16
45	181	57.3	566	30	AA653017	nc189003.s1	NCI_CGAP_Pr	7.42e-16

## ALIGNMENTS

RESULT 1  
LOCUS AA228953 373 bp mRNA  
DEFINITION nc14604.r1 NCI\_CGAP.Pr1 Homo sapiens cDNA clone IMAGE:1008078  
similar to gb:M21895 PROSTATE SPECIFIC ANTIGEN PRECURSOR (HUMAN).  
ACCESSION AA228953  
VERSION AA228953.1 GI:1651772  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 373)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT On Sep 12, 1996 this sequence version replaced gi:1394945.  
Contact: Robert Strausberg, Ph.D.  
Tel: (301) 496-1550  
Email: Robert.Strausberg@nih.gov  
Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuang,  
M.D., Michael Emmert-Buck, M.D., Ph.D.

Statistics: Mean 53.419; Variance 70.507; scale 0.758  
Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No. Query Match Length DB ID Description Pred. No.

CDNA Library Preparation: David B. Krizman, Ph.D.  
 CDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 www-bio.llnl.gov/bdrrp/image/image.html

## FEATURES

## Source

Seq primer: -28m13 rev1 ET from Amersham  
 High quality sequence stop: 342.

## Location/Qualifiers

1. .373  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:1008078"  
 /clone\_lib="NCI-CGAP\_Prl"  
 /sex="Male"  
 /dev\_stage="45 years old"  
 /lab\_host="DH10B"  
 /note="Vector: PAMPI0; Site\_1: NotI; Site\_2: EcoRI; 1st  
 strand cDNA was primed with oligo(dT)17 on 50 ng of  
 DNase-treated, total cellular RNA obtained from  
 5,000-10,000 microdissected, histologically normal  
 prostate epithelial cells. Double-stranded cDNA was  
 ligated to EcoRI adaptors, 5 cycles of PCR applied to the  
 cDNA with an adaptor-specific primer, and the resulting  
 PCR product subcloned into PAMPI0 by the UDG-cloning  
 method (Life Technologies). Average insert size is 600  
 bp. NOTE: Not directionally cloned. This library was  
 constructed by David Krizman."  
 BASE COUNT 78 a 106 c 107 g 82 t  
 ORIGIN

Query Match 98.7%; Score 312; DB 24; Length 373;  
 Best Local Similarity 60.8%; Pred. No. 1,08e-49;  
 Matches 45; Conservative 19; Mismatches 10; Indels 0; Gaps 0;

Db 182 GAGGAGTCTTGAGCCCAAGAACTGAGTGTGACCTCAGTATTTCATGAC 241

Qy 1 GARGARTTYTACNCANARARATNCARCTGTCGATGATGATGATGATGATGAT 60

Db 242 GTGTGTGCGCAGT 255

Qy 61 GTNTGYGCNCARGT 74

RESULT 2 AA228822 388 bp mRNA EST 20-AUG-1997

LOCUS nci15a04.s1 NCI-CGAP\_Prl Homo sapiens CDNA clone IMAGE:1008174

DEFINITION similar to gb:M21895 PROSTATE SPECIFIC ANTIGEN PRECURSOR (HUMAN);

ACCESSION AA228822

VERSION AA228822.1 GI:1851679

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 388)

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Unpublished (1997)

JOURNAL On Sep 12, 1996 this sequence version replaced gi:1394773.

COMMENT Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

www-bio.llnl.gov/bdrrp/image/image.html

Insert Length: 648 Std Error: 0.00

Seq primer: -41m13 fwd. ET from Amersham

High quality sequence stop: 341.

## FEATURES

## Source

1. .388  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:1008174"  
 /clone\_lib="NCI-CGAP\_Prl"  
 /sex="Male"  
 /dev\_stage="45 years old"  
 /lab\_host="DH10B"  
 /note="Vector: PAMPI0; Site\_1: NotI; Site\_2: EcoRI; 1st  
 strand cDNA was primed with oligo(dT)17 on 50 ng of  
 DNase-treated, total cellular RNA obtained from  
 5,000-10,000 microdissected, histologically normal  
 prostate epithelial cells. Double-stranded cDNA was  
 ligated to EcoRI adaptors, 5 cycles of PCR applied to the  
 cDNA with an adaptor-specific primer, and the resulting  
 PCR product subcloned into PAMPI0 by the UDG-cloning  
 method (Life Technologies). Average insert size is 600  
 bp. NOTE: Not directionally cloned. This library was  
 constructed by David Krizman."  
 BASE COUNT 81 a 119 c 102 g 86 t  
 ORIGIN

Query Match 98.7%; Score 312; DB 24; Length 388;  
 Best Local Similarity 60.8%; Pred. No. 1,08e-49;  
 Matches 45; Conservative 19; Mismatches 10; Indels 0; Gaps 0;

Db 262 ACTTGCGCAGCAGCTGATTTGGAATATACAGAGTCCACACTGATTTCTTGGG 321

Cp 74 ACYGGNGCRCAACRATCTCTTNSWDATNACRTGNARCTCNACRCAVTGNARVTTTNGN 15

Db 322 GTCAAGACTCCTC 335

Cp 14 GTNARAAVTCYTC 1

RESULT 3 AA228836 420 bp mRNA EST 20-AUG-1997

LOCUS nci16h06.r1 NCI-CGAP\_Prl Homo sapiens CDNA clone IMAGE:1008347

DEFINITION similar to gb:M21895 PROSTATE SPECIFIC ANTIGEN PRECURSOR (HUMAN);

ACCESSION AA228836

VERSION AA228836.1 GI:1851723

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 420)

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Unpublished (1997)

JOURNAL On Sep 12, 1996 this sequence version replaced gi:1394901.

COMMENT Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

www-bio.llnl.gov/bdrrp/image/image.html

Insert Length: 1270 Std Error: 0.00

Seq primer: -28m13 rev1 ET from Amersham

High quality sequence stop: 341.



strand cDNA was primed with oligo(dT)17 on 50 ng of DNase-treated, total cellular RNA obtained from 5,000-10,000 microdissected cells histologically-determined to be fully malignant prostate cancer cells. Double-stranded cDNA was ligated to EcoRI adaptors, 5 cycles of PCR applied to the cDNA with an adaptor-specific primer, and the resulting PCR product subcloned into pAMP10 by the UDG-cloning method (Life Technologies). Average insert size is 600 bp. NOTE: Not directionally cloned. This library was constructed by David Krizman.

BASE COUNT 94 a 127 c 127 g 88 t

Query Match 98.7%; Score 312; DB 30; Length 436;  
Best Local Similarity 60.8%; Pred. No. 1.08e-49;  
Matches 45; Conservative 19; Mismatches 10; Indels 0; Gaps 0;

Db 138 GAGGAGTCTTGACCCCAAGAACTCAGTGTGACCTCATGTATTTCATGAC 197  
||:||||:| || ||:||||:| ||:||||:| ||:||||:| ||:||||:| ||:||||:|  
QY 1 GARGARTTYTNACNCNAAARATNCARTGTGTCATGATTCATGATTCATGAT 60

Db 198 GTGTGTGCGCAAGT 211  
||:||||:| || ||:||||:| ||:||||:| ||:||||:| ||:||||:| ||:||||:|  
QY 61 GTNTGYGCNCARGT 74

RESULT 6  
LOCUS AA569482 439 bp mRNA EST 25-AUG-1997  
DEFINITION nt23a11.s1 NCI-CGAP\_Pr1 Homo sapiens cDNA clone IMAGE:914588  
similar to gb:M21895 PROSTATE SPECIFIC ANTIGEN PRECURSOR (HUMAN);,  
mRNA sequence.

ACCESSION AA569482  
VERSION AA569482.1 GI:2343462  
KEYWORDS EST.

SOURCE human.  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 439)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index

JOURNAL Unpublished (1997)  
COMMENT On Sep 12, 1996 this sequence version replaced gi:1407230.  
Contact: Robert Strausberg, Ph.D.

Email: Robert.Strausberg@nih.gov  
Tel: (301) 496-1550  
Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuagui,  
M.D., Michael Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: David B. Krizman, Ph.D.  
CDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bdrp/image/image.html](http://www-bio.llnl.gov/bdrp/image/image.html)

FEATURES  
source  
Seg primer: -40m13 fwd. ET from Amersham  
High quality sequence stop: 427.  
Location/Qualifiers  
1..439

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_image="914588"  
/clone\_lib="NCI-CGAP\_Pr1"  
/sex="Male"  
/dev\_stage="45 years old"  
/lab\_host="DH10B"  
/note="Vector: pAMP10; Site\_1: NotI; Site\_2: EcoRI; 1st  
strand cDNA was primed with oligo(dT)17 on 50 ng of  
DNase-treated, total cellular RNA obtained from  
5,000-10,000 microdissected, histologically normal  
prostate epithelial cells. Double-stranded cDNA was

ligated to EcoRI adaptors, 5 cycles of PCR applied to the cDNA with an adaptor-specific primer, and the resulting PCR product subcloned into pAMP10 by the UDG-cloning method (Life Technologies). Average insert size is 600 bp. NOTE: Not directionally cloned. This library was constructed by David Krizman.

BASE COUNT 94 a 128 c 124 g 93 t

Query Match 98.7%; Score 312; DB 29; Length 439;  
Best Local Similarity 60.8%; Pred. No. 1.08e-49;  
Matches 45; Conservative 19; Mismatches 10; Indels 0; Gaps 0;

Db 211 GAGGAGTCTTGACCCCAAGAACTCAGTGTGACCTCATGTATTTCATGAC 270  
||:||||:| || ||:||||:| ||:||||:| ||:||||:| ||:||||:| ||:||||:|  
QY 1 GARGARTTYTNACNCNAAARATNCARTGTGTCATGATTCATGATTCATGAT 60

Db 271 GTGTGTGCGCAAGT 284  
||:||||:| || ||:||||:| ||:||||:| ||:||||:| ||:||||:| ||:||||:|  
QY 61 GTNTGYGCNCARGT 74

RESULT 7  
LOCUS AA573727 453 bp mRNA EST 12-SEP-1997  
DEFINITION nt43e09.s1 NCI-CGAP\_Pr2 Homo sapiens cDNA clone IMAGE:916552  
similar to gb:M21895 PROSTATE SPECIFIC ANTIGEN PRECURSOR (HUMAN);,  
mRNA sequence.

ACCESSION AA573727  
VERSION AA573727.1 GI:2348242  
KEYWORDS EST.

SOURCE human.  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 453)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index

JOURNAL Unpublished (1997)  
COMMENT On Dec 30, 1996 this sequence version replaced gi:1528972.  
Contact: Robert Strausberg, Ph.D.

Email: Robert.Strausberg@nih.gov  
Tel: (301) 496-1550  
Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuagui,  
M.D., Michael Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: David B. Krizman, Ph.D.  
CDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bdrp/image/image.html](http://www-bio.llnl.gov/bdrp/image/image.html)

Insert Length: 1098 Std Error: 0.00  
Seg primer: -40m13 fwd. ET from Amersham  
High quality sequence stop: 387.  
Location/Qualifiers  
1..453

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_image="916552"  
/clone\_lib="NCI-CGAP\_Pr2"  
/sex="Male"  
/dev\_stage="45 years old"  
/lab\_host="DH10B"  
/note="Vector: pAMP10; Site\_1: NotI; Site\_2: EcoRI; 1st  
strand cDNA was primed with oligo(dT)17 on 50 ng of  
DNase-treated, total cellular RNA obtained from  
5,000-10,000 microdissected preneoplastic cells  
histologically-determined to be prostatic intraepithelial  
neoplasia 2 (PIN2) cells. Double-stranded cDNA was  
ligated to EcoRI adaptors, 5 cycles of PCR applied to the  
cDNA with an adaptor-specific primer, and the resulting  
PCR product subcloned into pAMP10 by the UDG-cloning





[illegible]

Db	231	GTGCTGGCGCAAGT	244
Oy	61	GTNTGTGCNCARGT	74
RESULT	11		
LOCUS	AA658261	489 bp	mRNA
DEFINITION	nuc12c03.s1 NCI CGAP_Pr2 Homo sapiens CDNA clone IMAGE:1208644 similar to gp:M1895 PROSTATE SPECIFIC ANTIGEN PRECURSOR (HUMAN);,		
ACCESSION	AA658261		
VERSION	AA658261.1	GI:2594415	
SOURCE	EST.		
ORGANISM	human.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;		
AUTHORS	Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.		
JOURNAL	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),		
COMMENT	Tumor Gene Index		
	Unpublished (1997)		
	On May 8, 1995 this sequence version replaced gi:801261.		
	Contact: Robert Strausberg, Ph.D.		
	Tel.: (301) 496-1350		
	Email: Robert.Strausberg@nih.gov		
	Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuangli,		
	M.D., Michael Emmert-Buck, M.D., Ph.D.		
	CDNA Library Preparation: David B. Krizman, Ph.D.		
	CNA Library Arrayed by: genome Systems Inc., Greg Lennon, Ph.D.		
	DNA Sequencing by: Washington University Genome Sequencing Center		
	clone distribution: NCI-CGAP clone distribution information can be		
	found through the I.W.A.G.E. Consortium/LNLT at:		
	www.bto.llnl.gov/dbt/p/image/image.html		
FEATURES			
source	Seq primer: -40m13 fwd. ET from Amersham		
	High quality sequence stop: 367.		
	Location/Qualifiers		
	1..489		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone_image="1208644"		
	/clone_lib="NCI_CGAP_Pr2"		
	/sex="Male"		
	/dev_stage="45 years old"		
	/lab_host="DH10B"		
	/note="Vector: PAMPI0; Site_1: NotI; Site_2: EcoRI; 1st		
	strand cDNA was primed with oligo(dT)17 on 50 ng of		
	DNAse-treated, total cellular RNA obtained from		
	5,000-10,000 microdissected preneoplastic cells		
	histologically-determined to be prostatic interepithelial		
	neoplasia 2 (PIN2) cells. Double-stranded cDNA was		
	ligated to EcoRI adaptors. 5 cycles of PCR applied to the		
	cDNA with an adaptor-specific primer, and the resulting		
	PCR product subcloned into PAMPI0 by the UDP-cloning		
	method (Life Technologies). Average insert size is 600		
	bp. NOTE: Not directionally cloned. This library was		
	constructed by David Krizman."		
BASE COUNT	119 a	117 c	138 g
ORIGIN		115 t	
Query Match	98.7%;	Score 312;	DB 30;
Best Local Similarity	60.8%;	Pred. No. 1.08e-49;	Length 489;
Matches	45;	Conservative 19;	Mismatches 10;
		Indels 0;	Gaps 0;
Db	15	GAGGACTTGTCAGCCCAAGAACAATCATGTGTGGACCTCAGTTATTTCGAATGAC	74
Oy	1	GAGGARTTYTNACCMAAARAARATYTCATGTGTGTAATTCATGATWATHWSAAATGAY	60
Db	75	GTCTGTGGCGCAAGT	88
Oy	61	GTNTGTGCNCARGT	74

RESULT 12  
LOCUS AA225115 491 bp mRNA EST 19-AUG-1997  
DEFINITION nc21g11.s1 NCI\_CGAP\_P11 Homo sapiens CDNA clone IMAGE:100836  
similar to gb:M21895 PROSTATE SPECIFIC ANTIGEN PRECURSOR (HUMAN);,  
mRNA sequence.

ACCESSION AA225115  
VERSION AA225115.1 GI:1846489  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 491)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
On Apr 14, 1993 this sequence version replaced gi:692766.  
COMMENT Contact: Robert Strausberg, Ph.D.  
Tel: (301) 496-1550  
Email: Robert.Strausberg@nih.gov  
Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuquai,  
M.D., Michael Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: David B. Krizman, Ph.D.  
CDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www.bio.llnl.gov/bdrp/image/image.html](http://www.bio.llnl.gov/bdrp/image/image.html)

JOURNAL  
On Apr 14, 1993 this sequence version replaced gi:692766.  
COMMENT Contact: Robert Strausberg, Ph.D.  
Tel: (301) 496-1550  
Email: Robert.Strausberg@nih.gov  
Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuquai,  
M.D., Michael Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: David B. Krizman, Ph.D.  
CDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www.bio.llnl.gov/bdrp/image/image.html](http://www.bio.llnl.gov/bdrp/image/image.html)

FEATURES  
source  
Insert Length: 1127 Std Error: 0.00  
Seq primer: -41m13 fwd. ET from Amersham  
High quality sequence stop: 335.  
Location/Qualifiers  
1..491  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_image="100836"  
/clone\_11b="NCI\_CGAP\_P11"  
/sex="Male"  
/dev\_stage="45 years old"  
/lab\_host="DH10B"  
/note="Vector: PAMP10; Site\_1: NotI; Site\_2: EcoRI; 1st  
strand cDNA was primed with oligo(dT)17 on 50 ng of  
DNase-treated, total cellular RNA obtained from  
5,000-10,000 microdissected, histologically normal  
prostate epithelial cells. Double-stranded cDNA was  
ligated to EcoRI adaptors, 5 cycles of PCR applied to the  
cDNA with an adaptor-specific primer, and the resulting  
PCR product subcloned into PAMP10 by the UDG-cloning  
method (Life Technologies). Average insert size is 600  
bp. NOTE: Not directionally cloned. This library was  
constructed by David Krizman."

BASE COUNT 121 a 128 c 128 g 111 t 3 others  
ORIGIN  
Query Match 98.7%; Score 312; DB 24; Length 491;  
Best Local Similarity 60.8%; Pred. No. 1,08e-49;  
Matches 45; Conservative 19; Mismatches 10; Indels 0; Gaps 0;

Db 55 GAGGAGTCTTGACCCCAAGAACTCAGTGTGACCTCATGTTATTCATGAC 114  
|||:||||:| |||:||||:| |||:||||:| |||:||||:| |||:||||:|  
QY 1 GARGARTTYTNACNCNAAARAATYTCAGTGTGATGATGATGATGATGATGAT 60  
|||:||||:| |||:||||:| |||:||||:| |||:||||:| |||:||||:|

Db 115 GTGTGTGCGCAAGT 128  
|||:||||:| |||:||||:| |||:||||:| |||:||||:| |||:||||:|  
QY 61 GTNTGYGCNCARGT 74  
|||:||||:| |||:||||:| |||:||||:| |||:||||:| |||:||||:|

RESULT 13  
LOCUS AA574023 505 bp mRNA EST 12-SEP-1997

DEFINITION  
n144f11.s1 NCI\_CGAP\_P12 Homo sapiens CDNA clone IMAGE:916653  
similar to gb:M21895 PROSTATE SPECIFIC ANTIGEN PRECURSOR (HUMAN);,  
mRNA sequence.

ACCESSION AA574023  
VERSION AA574023.1 GI:2348538  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 505)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
On Apr 14, 1993 this sequence version replaced gi:692976.  
COMMENT Contact: Robert Strausberg, Ph.D.  
Tel: (301) 496-1550  
Email: Robert.Strausberg@nih.gov  
Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuquai,  
M.D., Michael Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: David B. Krizman, Ph.D.  
CDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www.bio.llnl.gov/bdrp/image/image.html](http://www.bio.llnl.gov/bdrp/image/image.html)

JOURNAL  
On Apr 14, 1993 this sequence version replaced gi:692976.  
COMMENT Contact: Robert Strausberg, Ph.D.  
Tel: (301) 496-1550  
Email: Robert.Strausberg@nih.gov  
Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuquai,  
M.D., Michael Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: David B. Krizman, Ph.D.  
CDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www.bio.llnl.gov/bdrp/image/image.html](http://www.bio.llnl.gov/bdrp/image/image.html)

FEATURES  
source  
Insert Length: 1083 Std Error: 0.00  
Seq primer: -40m13 fwd. ET from Amersham  
High quality sequence stop: 348.  
Location/Qualifiers  
1..505  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_image="916653"  
/clone\_11b="NCI\_CGAP\_P12"  
/sex="Male"  
/dev\_stage="45 years old"  
/lab\_host="DH10B"  
/note="Vector: PAMP10; Site\_1: NotI; Site\_2: EcoRI; 1st  
strand cDNA was primed with oligo(dT)17 on 50 ng of  
DNase-treated, total cellular RNA obtained from  
5,000-10,000 microdissected preneoplastic cells  
histologically-determined to be prostatic intraepithelial  
neoplasia 2 (PIN2) cells. Double-stranded cDNA was  
ligated to EcoRI adaptors, 5 cycles of PCR applied to the  
cDNA with an adaptor-specific primer, and the resulting  
PCR product subcloned into PAMP10 by the UDG-cloning  
method (Life Technologies). Average insert size is 600  
bp. NOTE: Not directionally cloned. This library was  
constructed by David Krizman."

BASE COUNT 121 a 124 c 136 g 124 t  
ORIGIN  
Query Match 98.7%; Score 312; DB 29; Length 505;  
Best Local Similarity 60.8%; Pred. No. 1,08e-49;  
Matches 45; Conservative 19; Mismatches 10; Indels 0; Gaps 0;

Db 9 GAGGAGTCTTGACCCCAAGAACTCAGTGTGACCTCATGTTATTCATGAC 68  
|||:||||:| |||:||||:| |||:||||:| |||:||||:| |||:||||:|  
QY 1 GARGARTTYTNACNCNAAARAATYTCAGTGTGATGATGATGATGATGATGAT 60  
|||:||||:| |||:||||:| |||:||||:| |||:||||:| |||:||||:|

Db 69 GTGTGTGCGCAAGT 82  
|||:||||:| |||:||||:| |||:||||:| |||:||||:| |||:||||:|  
QY 61 GTNTGYGCNCARGT 74  
|||:||||:| |||:||||:| |||:||||:| |||:||||:| |||:||||:|

RESULT 14  
LOCUS AA534235 520 bp mRNA EST 21-AUG-1997  
DEFINITION n170e05.s1 NCI\_CGAP\_P10 Homo sapiens CDNA clone IMAGE:997856  
similar to gb:M21895 PROSTATE SPECIFIC ANTIGEN PRECURSOR (HUMAN);,  
mRNA sequence.

ACCESSION	AA534235	GI:2278251	
VERSION	AA534235.1		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;		
AUTHORS	Eutheria; Primates; Catarrhini; Homidae; Homo.		
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.		
JOURNAL	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),		
COMMENT	Tumor Gene Index		
	Unpublished (1997)		
	On May 8, 1995 this sequence version replaced gi 801205.		
	Contact: Robert Strausberg, Ph.D:		
	Tel: (301) 496-1550		
	Email: Robert.Strausberg@nih.gov		
	Tissue procurement: W. Marston Linehan, M.D., Rodrigo F. Chuquí,		
	M.D., Michael R. Emmert-Buck, M.D., Ph.D.		
	cDNA Library Preparation: David B. Kitzman, Ph.D.		
	DNA sequencing by: Washington University Genome Sequencing Center		
	Clone distribution: NCI-CGAP clone distribution information can be		
	found through the I.M.A.G.E. Consortium/ILNL at:		
	www.bio.lnl.gov/bdrp/imag/image.html		
FEATURES			
source	Insert Length: 1192 Std Error: 0.00		
	Seq primer: -40ml3 fwd. ET from Amersham		
	High quality sequence stop: 398.		
	Location/Qualifiers		
	1..520		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:997856"		
	/clone_lib="NCI-CGAP_Pri10"		
	/sex="male"		
	/tissue_type="invasive prostate tumor"		
	/lab_host="DH10B"		
	/note="Organ: prostate; Vector: pAMP10; mRNA made from		
	invasive prostate tumor, cDNA made by oligo-dT priming.		
	Non-directionally cloned. Size selected on agarose gel,		
	average insert size 600 bp. Library made by D. Kitzman,		
	NH."		
BASE COUNT	124 a 142 c 139 g 115 t		
ORIGIN			
Query Match	98.7%; Score 312; DB 28; Length 520;		
Best Local Similarity	60.8%; Pred. No.1.08e-49;		
Matches	45; Conservative 19; Mismatches 10; Indels 0; Gaps 0;		
Db	92 GAGGATTCCTGACCCCAAGAAACTCATGCTGACCTCCTGATTTCCANTGAC 151		
	:     :     :     :     :     :     :     :     :     :		
OY	1 GAGGATTTTNNCNCNMAARARNTNCARTGTGTGATTCATCACTNNATWENALYAY 60		
Db	152 GTGTGTGCCCACT 165		
	:     :     :     :     :     :     :     :     :     :		
OY	61 GINTGYGCNCARCT 74		
RESULT	15		
LOCUS	AA654548 542 bp mRNA EST 04-NOV-1997		
DEFINITION	nt59d11.s1 NCI-CGAP_Pri3 Homo sapiens cDNA clone IMAGE:1202805		
	similar to gb:U12895 PROSTATE SPECIFIC ANTIGEN PRECURSOR (HUMAN);		
	mRNA sequence.		
ACCESSION	AA654548		
VERSION	AA654548.1 GI:2590702		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;		
	Eutheria; Primates; Catarrhini; Homidae; Homo.		
REFERENCE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.		
AUTHORS	1 (bases 1 to 542)		
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),		

```

JOURNAL
Unpublished (1997)
On Sep 12, 1996 this sequence version replaced gi:1393439.
COMMENT
Tumor Gene Index
Email: Robert.Strausberg@nih.gov
Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuagui,
M.D., Michael Emmert-Buck, M.D., Ph.D.
CDNA library Preparation: David B. Krizman, Ph.D.
DNA Sequencing by: Genome Systems Inc., Greg Lennon, Ph.D.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLT at:
www-bio.lnlt.gov/dbfp/image/image.html

FEATURES
source
1. 542
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1202805"
/clone_1lb="NCI-CGAP_P13"
/sex="Male"
/dev_stage="45 years old"
/lab_host="DH10B"
/note="Vector: PAMP10; Site_1: NotI; Site_2: EcoRI; 1st
strand cDNA was primed with oligo(dT)17 on 50 ng of
DNase-treated, total cellular RNA obtained from
5,000-10,000 microdissected cells
histologically-determined to be fully malignant prostate
cancer cells. Double-stranded cDNA was ligated to EcoRI
adaptors, 5 cycles of PCR applied to the cDNA with an
adaptor-specific primer, and the resulting PCR product
subcloned into PAMP10 by the UDG-cloning method (life
technologies). Average insert size is 600 bp. NOTE: Not
directionally cloned. This library was constructed by
David Krizman."

BASE COUNT      119 a      137 c      167 g      119 t
ORIGIN
Query Match      98.7%: Score 312; DB 30; Length 542;
Best Local Similarity 60.8%: Prid. No. 1,08e-45;
Matches 45; Conservative 19; Mismatches 10; Indels 0; Gaps 0;

Db      103 GAGGAGTTCCTTGACCCCAAGAACTCGTGTGGACGCCATGTTATTCACATGAC 162
      |||:||||:| |||:||||:| |||:||||:| |||:| |||:| |||:| |||:|
Qy      1 GARGARTTYTNAACNANARARATYNCATGTYGNGATYTNCAATGATNATHWSNAATGAY 60
      |||:||||:| |||:||||:| |||:||||:| |||:| |||:| |||:| |||:|

Db      163 GTGTGTGGCAAGT 176
      |||:||||:| |||:||||:| |||:||||:| |||:| |||:| |||:| |||:|
Qy      61 GTNTGTGCGNCAGT 74

```

Search completed: Fri Jun 23 17:10:33 2000  
Job time : 4615 secs.

Size limit = 25 AA

WATERMAN

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

Mpsrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Thu Jun 22 21:27:01 2000; MasPar time 4.15 Seconds  
Tabular output not generated. 142,760 Million cell updates/sec

Title: >US-09-332-866-1  
Description: (1-25) from US09332866.pep  
Perfect Score: 163  
Sequence: 1 EEFLLPKKLCQVDLHVSNDVCAQV 25

Scoring table:  
PAM 150  
Gap 15

Searched: 188963 seqs, 23686106 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries  
Maximum DB seq length 25

Database: a-geneseq35  
1:geneseqp

Statistics: Mean 20.967; Variance 66.122; scale 0.317

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	116	63.4	15	1	WS8047 Human prostate specific	4.36e-05
2	111	60.7	15	1	WS8048 Human prostate specific	1.67e-04
3	110	60.1	15	1	WS8050 Human prostate specific	2.19e-04
4	108	59.0	15	1	WS8049 Human prostate specific	3.73e-04
5	98	53.6	18	1	RES193 Residues 140-157 of th	5.23e-03
6	96	52.5	20	1	W11022 Antigen peptide derive	8.82e-03
7	93	50.8	15	1	WS8046 Human prostate specific	1.92e-02
8	90	48.2	15	1	WS8051 Human prostate specific	4.16e-02
9	90	49.2	17	1	W30784 Prostate Specific Antl	4.16e-02

Note: Post-processor removed 36 summaries from list due to search parameters chosen.

## ALIGNMENTS

RESULT 1  
ID WS8047 standard; peptide; 15 AA.  
AC WS8047;  
DT 11-AUG-1998 (first entry)  
DE Human prostate specific antigen peptide SEQ ID NO:65.  
KW Human; prostate specific antigen; PSA; epitope; monoclonal antibody;  
detection; cancer; serine protease.

OS Synthetic.  
OS Homo sapiens.  
PN WS8048;  
PD 12-MAR-1998.  
PF 25-AUG-1997; Un4909.  
PR 06-SEP-1996; US-025404.  
PI (GEN2 ) CENTOCOR INC.  
PI Heavner GA;  
DR WPI: 98-193789/17.  
PT Monoclonal antibodies specific for prostate specific antigen -  
PT useful, e.g. in screening for prostate or breast cancer and  
PT especially to distinguish between benign prostatic hyperplasia and  
PT prostate cancer.  
PS Example: Page 56; 84pp; English.  
CC The present sequence represents a prostate specific antigen (PSA)  
CC peptide. Monoclonal antibodies specific for PSA and hybridomas producing  
CC them have been developed. The antibodies: (a) bind to free PSA, (b) are  
CC monoclonal antibody 365 binding to amino acids 82-87 of free PSA (LKNRFL)  
CC or fragment, and (c) are monoclonal antibodies 10, 11, 16 or 22.2 and  
CC 15.2, 156 or 225 binding to amino acids 139-144 (EEFLRP) and 55-60  
CC (SLFPE) respectively of free and bound PSA, or fragments. The antibodies  
CC are useful to detect PSA. For detecting free PSA only, an immunoassay  
CC comprising a solid support with attached free PSA only, an immunoassay  
CC specific for free PSA (especially (b)) and a PSA standard can be used.  
CC To detect both free and bound PSA, a second solid support with attached  
CC (differently labelled) monoclonal antibody binding free and bound PSA  
CC (especially selected from (c)) can be used either with, or in place of,  
CC the first solid support. The antibodies are useful in cancer screening,  
CC especially prostate and breast cancer. By obtaining total and free PSA  
CC values, their ratio can be used to separate prostatic cancer (PCA) from  
CC benign prostatic hyperplasia (BPH) patients. Measurement of PSA is also  
CC useful after radical prostatectomy, to predict disease persistence. The  
CC antibodies allow evaluation of PSA free/total ratio, enabling separation  
CC of BPH and PCA patients with PSA values 4-10 ng/ml not previously  
CC possible by total PSA testing, avoiding biopsies.  
SQ Sequence 15 AA:

Query Match 63.4%; Score 116; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 4.36e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 EEFLLPKKLCQVDLH 15  
0Y 1 EEFLLPKKLCQVDLH 15

RESULT 2  
ID WS8048 standard; peptide; 15 AA.  
AC WS8048;  
DT 11-AUG-1998 (first entry)  
DE Human prostate specific antigen peptide SEQ ID NO:66.  
KW Human; prostate specific antigen; PSA; epitope; monoclonal antibody;  
detection; cancer; serine protease.  
OS Synthetic.  
OS Homo sapiens.  
PN WS8048;  
PD 12-MAR-1998.  
PF 25-AUG-1997; Un4909.  
PR 06-SEP-1996; US-025404.  
PI (GEN2 ) CENTOCOR INC.  
PI Heavner GA;  
DR WPI: 98-193789/17.  
PT Monoclonal antibodies specific for prostate specific antigen -  
PT useful, e.g. in screening for prostate or breast cancer and  
PT especially to distinguish between benign prostatic hyperplasia and  
PT prostate cancer.  
PS Example: Page 57; 84pp; English.  
CC The present sequence represents a prostate specific antigen (PSA)  
CC peptide. Monoclonal antibodies specific for PSA and hybridomas producing  
CC them have been developed. The antibodies: (a) bind to free PSA, (b) are  
CC monoclonal antibody 365 binding to amino acids 82-87 of free PSA (LKNRFL)  
CC or fragment, and (c) are monoclonal antibodies 10, 11, 16 or 22.2 and  
CC 15.2, 156 or 225 binding to amino acids 139-144 (EEFLRP) and 55-60

CC (SLRPE) respectively of free and bound PSA, or fragments. The antibodies  
 CC are useful to detect PSA. For detecting free PSA only, an immunoassay  
 CC comprising a solid support with attached labelled monoclonal antibody  
 CC specific for free PSA (especially (b)) and a PSA standard can be used.  
 CC To detect both free and bound PSA, a second solid support with attached  
 CC (differently labelled) monoclonal antibody binding free and bound PSA  
 CC (especially selected from (c)) can be used either with, or in place of,  
 CC the first solid support. The antibodies are useful in cancer screening,  
 CC especially prostate and breast cancer. By obtaining total and free PSA  
 CC values, their ratio can be used to separate prostatic cancer (PCA) from  
 CC benign prostatic hyperplasia (BHP) patients. Measurement of PSA is also  
 CC useful after radical prostatectomy, to predict disease persistence. The  
 CC antibodies allow evaluation of PSA free/total ratio, enabling separation  
 CC of BPH and PCA patients with PSA values 4-10 ng/ml not previously  
 CC possible by total PSA testing, avoiding biopsies.

SO Sequence 15 AA;  
 Query Match 60.7%; Score 111; DB 1; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 1.67e-04;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 LTPKLCQVDLHVNTS 15  
 |||||  
 QY 4 LTPKLCQVDLHVNTS 18

RESULT 3  
 ID W58050 standard; peptide; 15 AA.

AC W58050.  
 DT 11-AUG-1998 (first entry)  
 DE Human prostate specific antigen peptide SEQ ID NO:68.  
 KW Human; prostate specific antigen; PSA; epitope; monoclonal antibody;  
 KM detection; cancer; serine protease.

OS Synthetic.  
 OS Homo sapiens.  
 PN W09810292-AL.  
 PD 12-MAR-1998.  
 PF 25-AUG-1997; U14909.  
 PR 06-SEP-1996; US-025404.  
 PA (GENZ ) CENTOCOR INC.  
 PI Heavner GA;

DR WPI: 98-193789/17  
 PT Monoclonal antibodies specific for prostate specific antigen -  
 PT useful, e.g. in screening for prostate or breast cancer and  
 PT especially to distinguish between benign prostatic hyperplasia and  
 PT prostate cancer  
 PS Example: Page 58; 84pp; English.

CC The present sequence represents a prostate specific antigen (PSA)  
 CC peptide. Monoclonal antibodies specific for PSA and hybridomas producing  
 CC them have been developed. The antibodies: (a) bind to free PSA; (b) are  
 CC monoclonal antibody 365 binding to amino acids 82-87 of free PSA (LKNRFL)  
 CC or fragment, and (c) are monoclonal antibodies 10, 11, 16 or 22-2 and  
 CC 15, 2, 156 or 225 binding to amino acids 139-144 (EELFRT) and 55-60  
 CC (SLRPE) respectively of free and bound PSA, or fragments. The antibodies  
 CC are useful to detect PSA. For detecting free PSA only, an immunoassay  
 CC comprising a solid support with attached labelled monoclonal antibody  
 CC specific for free PSA (especially (b)) and a PSA standard can be used.  
 CC To detect both free and bound PSA, a second solid support with attached  
 CC (differently labelled) monoclonal antibody binding free and bound PSA  
 CC (especially selected from (c)) can be used either with, or in place of,  
 CC the first solid support. The antibodies are useful in cancer screening,  
 CC especially prostate and breast cancer. By obtaining total and free PSA  
 CC values, their ratio can be used to separate prostatic cancer (PCA) from  
 CC benign prostatic hyperplasia (BHP) patients. Measurement of PSA is also  
 CC useful after radical prostatectomy, to predict disease persistence. The  
 CC antibodies allow evaluation of PSA free/total ratio, enabling separation  
 CC of BPH and PCA patients with PSA values 4-10 ng/ml not previously  
 CC possible by total PSA testing, avoiding biopsies.

SO Sequence 15 AA;  
 Query Match 60.1%; Score 110; DB 1; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 2.19e-04;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 QCVDLHVNSDYCAQ 15  
 |||||  
 QY 10 QCVDLHVNSDYCAQ 24

RESULT 4  
 ID W58049 standard; peptide; 15 AA.

AC W58049.  
 DT 11-AUG-1998 (first entry)  
 DE Human prostate specific antigen peptide SEQ ID NO:67.  
 KW Human; prostate specific antigen; PSA; epitope; monoclonal antibody;  
 KM detection; cancer; serine protease.

OS Synthetic.  
 OS Homo sapiens.  
 PN W09810292-AL.  
 PD 12-MAR-1998.  
 PF 25-AUG-1997; U14909.  
 PR 06-SEP-1996; US-025404.  
 PA (GENZ ) CENTOCOR INC.  
 PI Heavner GA;

DR WPI: 98-193789/17  
 PT Monoclonal antibodies specific for prostate specific antigen -  
 PT useful, e.g. in screening for prostate or breast cancer and  
 PT especially to distinguish between benign prostatic hyperplasia and  
 PT prostate cancer  
 PS Example: Page 57; 84pp; English.

CC The present sequence represents a prostate specific antigen (PSA)  
 CC peptide. Monoclonal antibodies specific for PSA and hybridomas producing  
 CC them have been developed. The antibodies: (a) bind to free PSA; (b) are  
 CC monoclonal antibody 365 binding to amino acids 82-87 of free PSA (LKNRFL)  
 CC or fragment, and (c) are monoclonal antibodies 10, 11, 16 or 22-2 and  
 CC 15, 2, 156 or 225 binding to amino acids 139-144 (EELFRT) and 55-60  
 CC (SLRPE) respectively of free and bound PSA, or fragments. The antibodies  
 CC are useful to detect PSA. For detecting free PSA only, an immunoassay  
 CC comprising a solid support with attached labelled monoclonal antibody  
 CC specific for free PSA (especially (b)) and a PSA standard can be used.  
 CC To detect both free and bound PSA, a second solid support with attached  
 CC (differently labelled) monoclonal antibody binding free and bound PSA  
 CC (especially selected from (c)) can be used either with, or in place of,  
 CC the first solid support. The antibodies are useful in cancer screening,  
 CC especially prostate and breast cancer. By obtaining total and free PSA  
 CC values, their ratio can be used to separate prostatic cancer (PCA) from  
 CC benign prostatic hyperplasia (BHP) patients. Measurement of PSA is also  
 CC useful after radical prostatectomy, to predict disease persistence. The  
 CC antibodies allow evaluation of PSA free/total ratio, enabling separation  
 CC of BPH and PCA patients with PSA values 4-10 ng/ml not previously  
 CC possible by total PSA testing, avoiding biopsies.

SO Sequence 15 AA;  
 Query Match 59.0%; Score 108; DB 1; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 3.73e-04;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 KRLQCVDLHVNSDV 15  
 |||||  
 QY 7 KRLQCVDLHVNSDV 21

RESULT 5  
 ID W58193 standard; peptide; 18 AA.

AC W58193.  
 DT 24-AUG-1995 (first entry)  
 DE Residues 140-157 of the deduced sequence of human kallikrein hK2.  
 KW Prostate specific glandular kallikrein; hK2; prostate cancer.

OS Homo sapiens.  
 OS Homo sapiens.  
 PN W09503334-A.  
 PD 02-FEB-1995.  
 PF 28-JUN-1994; U07329.  
 PR 22-JUL-1993; US-096946.  
 PA (MAYO-) MAYO FOUNDATION.  
 PI Klee GG, Tindall DJ, Young CYF, Young CY;  
 DR WPI: 95-075194/10.

PT New antibodies specific for human prostate glandular kallikrein  
 used for purifications and in assays which can be used for the  
 PT detection and staging of prostatic cancer  
 PS disclosure: Page 23; 43pp: English.  
 CC A polypeptide of the formula in R65190 is claimed. It represents  
 CC residues 41-56 of the deduced AA sequence of human prostate specific  
 CC glandular kallikrein (hk2) as reported by L.J. Schedlish et al., DNA,  
 CC 6, 429, 1987. Immunogenic or immunoreactive subunits of this peptide,  
 CC pref. of at least peptidyl units, which can be specifically bound  
 CC by specific anti-hk2 antibodies are also within the scope of the  
 CC invention. Other immunogenic hk2 subunits will also generate an  
 CC antibody response that is unique to hk2, including the sequences in  
 CC R65191-R65198. The numbering of the AAs identify the hk2 residues  
 CC disclosed in Schedlish et al., cited above.  
 SQ Sequence 18 AA;

DB Query Match 53.6%; Score 96; DB 1; Length 18;  
 Best Local Similarity 66.7%; Pred. NO. 5.23e-03;  
 Matches 12; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 EFLRPSRLOCVSLHLN 18 ✓  
 2 EFLRPSRLOCVSLHLN 19

8/3

RESULT 6  
 ID W11022 standard; peptide: 20 AA.  
 AC W11022;  
 DT 27-OCT-1997 (first entry)  
 DE Antigen peptide derived from human prostate specific antigen.  
 KW PSA; kallikrein: prostate cancer; antigenic; polyclonal antisera.  
 OS Homo sapiens.  
 PN W09640754-A1.  
 PD 19-DEC-1996.  
 PE 06-JUN-1996; U09303.  
 PR 07-JUN-1995; US-472228.  
 PA (TEXA ) UNIV TEXAS SYSTEM.  
 PI Fritsche HA, Johnston DA, Kokolus WJ;  
 DR WPI; 97-108633/10.  
 PT Antigen peptide derived from prostate specific antigen - does not  
 PT cross react with related kallikreins, for diagnosis of prostate  
 PT cancer

PS Example 2; Page 42; 74pp: English.  
 CC The present sequence represents a novel peptide which has 20 contiguous  
 CC amino acids derived from the 240 residue sequence of the human prostate  
 CC specific antigen (PSA) (see W11023). The preferred peptide has two  
 CC hydrophobic regions and one hydrophilic region each of about 5 aa in  
 CC length, arranged as follows: hydrophobic-hydrophilic-hydrophobic.  
 CC Polyclonal antisera specific for the antigenic peptide may be used in a  
 CC method for diagnosing prostate cancer in vivo or in vitro. The peptide  
 CC represents a sequence unique to PSA which does not cross react with  
 CC certain kallikreins.  
 SQ Sequence 20 AA;

DB Query Match 52.5%; Score 96; DB 1; Length 20;  
 Best Local Similarity 92.9%; Pred. NO. 8.81e-03;  
 Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 7 KRLOCVLHVHNSND 20 ✓  
 7 KRLOCVLHVHNSND 20

RESULT 7  
 ID W58046 standard; peptide: 15 AA.  
 AC W58046;  
 DT 11-AUG-1998 (first entry)  
 DE Human prostate specific antigen peptide SEQ ID NO:64.  
 KW Human; prostate specific antigen; PSA; epitope; monoclonal antibody;  
 KM detection; cancer; serine protease.  
 OS Synthetic.  
 PN W09810292-A1.

PD 12-MAR-1998.  
 PR 25-AUG-1997; U14909.  
 PR 06-SEP-1996; US-025404.  
 PA (CENZ ) CENTOCOR INC.  
 PI Heavner GA;  
 DR WPI; 98-193789/17.  
 PT Monoclonal antibodies specific for prostate specific antigen -  
 PT useful, e.g. in screening for prostate or breast cancer and  
 PT especially to distinguish between benign prostatic hyperplasia and  
 PT prostate cancer

PS Example; Page 36; 84pp: English.  
 CC The present sequence represents a prostate specific antigen (PSA)  
 CC peptide. Monoclonal antibodies specific for PSA and hydrolases producing  
 CC them have been developed. The antibodies: (a) bind to free PSA; (b) are  
 CC monoclonal antibody 365 binding to amino acids 82-87 of free PSA (LKNRFU)  
 CC or fragment, and (c) are monoclonal antibodies 10, 11, 16 or 22.2 and  
 CC 15.2, 156 or 225 binding to amino acids 139-144 (EFLRTP) and 55-60  
 CC (SLRHPD) respectively of free and bound PSA, or fragments. The antibodies  
 CC are useful to detect PSA. For detecting free PSA only, an immunoassay  
 CC comprising a solid support with attached labelled monoclonal antibody  
 CC specific for free PSA (especially (b)) and a PSA standard can be used.  
 CC To detect both free and bound PSA, a second solid support with attached  
 CC (differently labelled) monoclonal antibody binding free and bound PSA  
 CC (especially selected from (c)) can be used either with, or in place of,  
 CC the first solid support. The antibodies are useful in cancer screening,  
 CC especially prostate and breast cancer. By obtaining total and free PSA  
 CC values, their ratio can be used to separate prostatic cancer (PCA) from  
 CC benign prostatic hyperplasia (BHP) patients. Measurement of PSA is also  
 CC useful after radical prostatectomy, to predict disease persistence. The  
 CC antibodies allow evaluation of PSA free/total ratio, enabling separation  
 CC of BPH and PCA patients with PSA values 4-10 ng/ml not previously  
 CC possible by total PSA testing, avoiding biopsies.  
 SQ Sequence 15 AA;

DB Query Match 50.8%; Score 93; DB 1; Length 15;  
 Best Local Similarity 100.0%; Pred. NO. 1.92e-02;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EFLRPSRLOCV 12

RESULT 8  
 ID W58051 standard; peptide: 15 AA.  
 AC W58051;  
 DT 11-AUG-1998 (first entry)  
 DE Human prostate specific antigen peptide SEQ ID NO:69.  
 KW Human; prostate specific antigen; PSA; epitope; monoclonal antibody;  
 KM detection; cancer; serine protease.  
 OS Synthetic.  
 PN W09810292-A1.  
 PD 12-MAR-1998.  
 PR 25-AUG-1997; U14909.  
 PR 06-SEP-1996; US-025404.  
 PA (CENZ ) CENTOCOR INC.  
 PI Heavner GA;  
 DR WPI; 98-193789/17.  
 PT Monoclonal antibodies specific for prostate specific antigen -  
 PT useful, e.g. in screening for prostate or breast cancer and  
 PT especially to distinguish between benign prostatic hyperplasia and  
 PT prostate cancer

PS Example; Page 58; 84pp: English.  
 CC The present sequence represents a prostate specific antigen (PSA)  
 CC peptide. Monoclonal antibodies specific for PSA and hydrolases producing  
 CC them have been developed. The antibodies: (a) bind to free PSA; (b) are  
 CC monoclonal antibody 365 binding to amino acids 82-87 of free PSA (LKNRFU)  
 CC or fragment, and (c) are monoclonal antibodies 10, 11, 16 or 22.2 and  
 CC 15.2, 156 or 225 binding to amino acids 139-144 (EFLRTP) and 55-60  
 CC (SLRHPD) respectively of free and bound PSA, or fragments. The antibodies  
 CC are useful to detect PSA. For detecting free PSA only, an immunoassay  
 CC comprising a solid support with attached labelled monoclonal antibody



CC specific for free PSA (especially (b)) and a PSA standard can be used.  
 CC To detect both free and bound PSA, a second solid support with attached  
 CC (differently labeled) monoclonal antibody binding free and bound PSA  
 CC (especially selected from (c)) can be used either with, or in place of,  
 CC the first solid support. The antibodies are useful in cancer screening,  
 CC especially prostate and breast cancer. By obtaining total and free PSA  
 CC values, their ratio can be used to separate prostatic cancer (PCA) from  
 CC benign prostatic hyperplasia (BPH) patients. Measurement of PSA is also  
 CC useful after radical prostatectomy, to predict disease persistence. The  
 CC antibodies allow evaluation of PSA free/total ratio, enabling separation  
 CC of BPH and PCA patients with PSA values 4-10 ng/ml not previously  
 CC possible by total PSA testing, avoiding biopsies.  
 SQ Sequence 15 AA;

Query Match 49.2%; Score 90; DB 1; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 4.16e-02;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 DLHVISNDVCAQV 13  
 |||||  
 QY 13 DLHVISNDVCAQV 25

RESULT 9  
 ID W30784 standard; peptide: 17 AA.

AC W30784;  
 DT 27-FEB-1998 (first entry)  
 DE Prostate Specific Antigen (PSA) derived peptide ABR16.  
 KW Prostate specific antigen; PSA; immunogenic; human glandular kallikrein;  
 KW HK2; antigen; antibody; detection; diagnosis; prostate cancer.  
 OS Synthetic.  
 OS Homo sapiens.  
 PN MO9729199-A2.  
 PD 14-AUG-1997.  
 PE 06-FEB-1997; U01911.  
 PR 06-FEB-1996; US-595945.  
 PA (ABBO) ABBOTT LAB  
 PI Bridon DP, Dowell BL, Lilja H, Pettersson IK, Pihonen T;  
 PI Qiu X, Vihtinen MA;  
 DR WPI: 97-415352/38.  
 PT Prostate Specific Antigen peptide(s) - useful for diagnosis of  
 PT prostate cancer  
 PS Claim 2; Page 10; 42pp; English.  
 CC W30769-84 are synthetic peptides derived from the prostate specific  
 CC antigen (PSA) sequence. These peptides are identical to a highly  
 CC immunogenic region of PSA, and also comprise one or more amino acids  
 CC identical or non-identical to the amino acid sequence of human glandular  
 CC kallikrein (HK2). The peptides are used as antigens for the production of  
 CC antibodies which are used to detect PSA in a test sample (claimed). This  
 CC is useful for diagnosis of prostate cancer. The peptides enable the  
 CC production of antisera necessary to determine the amount of total PSA,  
 CC free PSA and PSA-ACP complex present in a sample and thus improve the  
 CC ability of the clinician to distinguish, e.g., between BPH (benign  
 CC prostatic hyperplasia) and prostatic cancer in a patient.  
 SQ Sequence 17 AA;

Query Match 49.2%; Score 90; DB 1; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 4.16e-02;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 DLHVISNDVCAQV 13  
 |||||  
 QY 13 DLHVISNDVCAQV 25

Search completed: Thu Jun 22 21:27:06 2000  
 Job time : 5 secs.



\*\*\*\*\*  
 M P E R E H  
 \*\*\*\*\*  
 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

Mpsrch\_p protein - protein database search, using Smith-Waterman algorithm  
 Run on: Thu Jun 22 21:24:52 2000; Maspar time 3.43 Seconds  
 Tabular output not generated. 105.343 Million cell updates/sec

Title: >US-09-332-866-1  
 Description: (1-25) from US09332866.pep  
 Perfect Score: 183  
 Sequence: 1 EFLTPKRLQCVDLHVISNDVCAQV 25

Scoring table:  
 PAM 150  
 Gap 15

Searched: 145341 seqs, 14437480 residues  
 Post-processing: Minimum Match 08  
 Listing first 45 summaries  
 Maximum DB seq length 25

Database: a-issued  
 1:5A\_COMB 2:5B\_COMB 3:6\_COMB 4:PCT\_COMB 5:backfiles1

Statistics: Mean 19.756; Variance 64.432; scale 0.307  
 Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Match	Length	DB	ID	Description	Pred. No.
1	98	53.6	18	4	PCT-US94-0	Sequence 4, Applicatio	3.47e-03	
2	98	53.6	18	1	US-08-096-	Sequence 4, Applicatio	3.47e-03	
3	96	52.5	20	1	US-08-472-	Sequence 22, Applicati	5.79e-03	
4	96	52.5	20	4	PCT-US96-0	Sequence 22, Applicati	5.79e-03	

Note: Post-processor removed 41 summaries from list due to search parameters chosen.

ALIGNMENTS

RESULT 1  
 ID PCT-US94-07329-4 STANDARD: PRT: 18 AA.  
 AC xxxxxx  
 XX  
 XX  
 DT  
 XX  
 DE Sequence 4, Application PC/TUS9407329  
 CC Sequence 4, Application PC/TUS9407329  
 CC GENERAL INFORMATION:  
 CC APPLICANT: Mayo Foundation for Medical

CC APPLICANT: Education and Research  
 CC TITLE OF INVENTION: Antibodies Specific for Human  
 CC TITLE OF INVENTION: Prostate Glandular kallikrein  
 CC NUMBER OF SEQUENCES: 11  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESS: Schweigman, Lundberg & Moessner  
 CC STREET: 3500 IDS Center  
 CC CITY: Minneapolis  
 CC STATE: MN  
 CC COUNTRY: USA  
 CC ZIP: 55402  
 CC  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Floppy disk  
 CC COMPUTER: IBM PC compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: Patent Release #1.0, Version #1.25  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: PCT/US94/07329  
 CC FILING DATE:  
 CC CLASSIFICATION:  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Raasch, Kevin W.  
 CC REGISTRATION NUMBER: 35,651  
 CC REFERENCE/DOCKET NUMBER: 150.62MO  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: 612-339-0331  
 CC TELEFAX: 612-339-3061  
 CC INFORMATION FOR SEQ ID NO: 4:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 18 amino acids  
 CC TYPE: amino acid  
 CC STRANDEDNESS: single  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: peptide  
 CC IMMEDIATE SOURCE:  
 CC CLONE: hK2 amino acid residues 140-157  
 CC SEQUENCE 18 AA: 2112 MW: 1815 CN;  
 CC  
 CC Query Match 53.6%; Score 98; DB 4; Length 18;  
 CC Best Local Similarity 66.7%; Pred. No. 3.47e-03;  
 CC Matches 12; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
 CC  
 CC Db 1 EFLRPSRLQCVSHLSLN 18  
 CC ||| | : ||| | : ||  
 CC Qy 2 EFLTPKRLQCVDLHVISN 19  
 CC  
 CC RESULT 2  
 CC ID US-08-096-946-4 STANDARD: PRT: 18 AA.  
 CC AC xxxxxx  
 CC XX  
 CC XX  
 CC DT  
 CC XX  
 CC DE Sequence 4, Application US/08096946  
 CC Patent No. 5516639  
 CC GENERAL INFORMATION:  
 CC APPLICANT: Tindall, Donald J.  
 CC APPLICANT: Young, Charles Y-F  
 CC APPLICANT: Klee, George G.  
 CC TITLE OF INVENTION: Antibodies Specific for Human  
 CC TITLE OF INVENTION: Prostate Glandular kallikrein  
 CC NUMBER OF SEQUENCES: 11  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESS: Merchant & Gould  
 CC STREET: 3100 No. 5516639west Center  
 CC CITY: Minneapolis  
 CC STATE: MN  
 CC COUNTRY: USA  
 CC ZIP: 55402

Mon Jun 26 14:52:31 2000

US-09-332-866-1a.rai

Page 2

CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patentin Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/096, 946  
CC FILING DATE: 19930722  
CC CLASSIFICATION: 436  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Woessner, Warren D.  
CC REGISTRATION NUMBER: 30,440  
CC REFERENCE/DOCKET NUMBER: 1016.62-US-01  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 612-332-5300  
CC TELEFAX: 612-332-8081  
CC INFORMATION FOR SEO ID NO: 4:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 18 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: peptide  
CC IMMEDIATE SOURCE:  
CC CLONE: HK2 amino acid residues 140-157  
CC SEQUENCE 18 AA; 2112 MW; 1815 CN;  
SQ

Query Match 53.6%; Score 98; DB 1; Length 18;  
Best Local Similarity 66.7%; Pred. No. 3,47e-03;  
Matches 12; Conservative 3; Mismatches 3; Indels 0; Gaps 0.

Db 1 EFLRPRSLQCVSLHLISN 18  
||| | : ||| | : ||| |  
QY 2 EFLTPRKLCQVDLHVYSN 19

RESULT 3 STANDARD: PRT; 20 AA.  
ID US-08-472-228A-22

XX xxxxxx  
XX  
XX  
DT  
XX  
DE Sequence 22, Application US/084722228A  
XX  
CC Sequence 22, Application US/084722228A  
CC Patent No. 5807978  
CC GENERAL INFORMATION:  
CC APPLICANT: Kokoius, William J.  
CC APPLICANT: Flitsche, Herbert A.  
CC APPLICANT: Johnston, Dennis A.  
CC TITLE OF INVENTION: IMMUNOGENIC PEPTIDES OF PROSTATE  
CC TITLE OF INVENTION: SPECIFIC ANTIGEN  
CC NUMBER OF SEQUENCES: 22  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Arnold, White & Durkee  
CC STREET: P. O. Box 4433  
CC CITY: Houston  
CC STATE: Texas  
CC COUNTRY: United States of America  
CC ZIP: 77210  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patentin Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/472,228A  
CC FILING DATE: 07-JUN-1995  
CC CLASSIFICATION: 424  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Highlander, Steven L.  
CC REGISTRATION NUMBER: 37,642

```

CC      REFERENCE/DOCKET NUMBER:  UTSC:388
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE: (512) 418-3000
CC      TELEFAX: (512) 474-7577
CC      TELEX: 79-0924
CC      INFORMATION FOR SEQ ID NO: 22:
CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH: 20 amino acids
CC      TYPE: amino acid
CC      STRANDEDNESS: single
CC      TOPOLOGY: linear
CC      MOLECULE TYPE: peptide
SQ      SEQUENCE 20 AA; 2389 MW; 2130 CN;

Query Match          52.5%; Score 96; DB 1; Length 20;
Best Local Similarity 92.9%; Pred. No. 5.79e-03;
Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db      7      KKLCQVOLHVISND 20
        |||||:|||||
OY      7      KKLCQVDLHVISND 20

RESULT 4
ID      PCT-US96-09303-22      STANDARD;      PRT;      20 AA.
XX      XXXXXX

Sequence 22, Application PC/TUS9609303
Sequence 22, Application PC/TUS9609303
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: IMMUNOGENIC PEPTIDES OF PROSTATE
TITLE OF INVENTION: SPECIFIC ANTIGEN
NUMBER OF SEQUENCES: 22
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/09303
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/472,228
FILING DATE: 07-JUN-1995
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
SQ      SEQUENCE 20 AA; 2389 MW; 2130 CN;

Query Match          52.5%; Score 96; DB 4; Length 20;
Best Local Similarity 92.9%; Pred. No. 5.79e-03;
Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db      7      KKLCQVOLHVISND 20
        |||||:|||||
OY      7      KKLCQVDLHVISND 20

```

\*\*\*\*\*  
 M P E S R E L H  
 (TM)  
 \*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

Msrch\_dp protein - protein database search, using Smith-Waterman algorithm  
 Run on: Thu Jun 22 21:26:38 2000; Maspar time 5.78 Seconds  
 Tabular output not generated. 203.949 Million cell updates/sec

Title: >US-09-332-866-1  
 Description: (1-25) from US09332866.pep  
 Perfect Score: 183  
 Sequence: 1 EEFLPKKLQCVDLHVISNDVCAQV 25

Scoring table:  
 PAM 150  
 Gap 15

Searched: 142080 seqs, 47172406 residues

Post-processing: Minimum Match 0%  
 Listing first 45 summaries  
 Maximum DB seq length 25

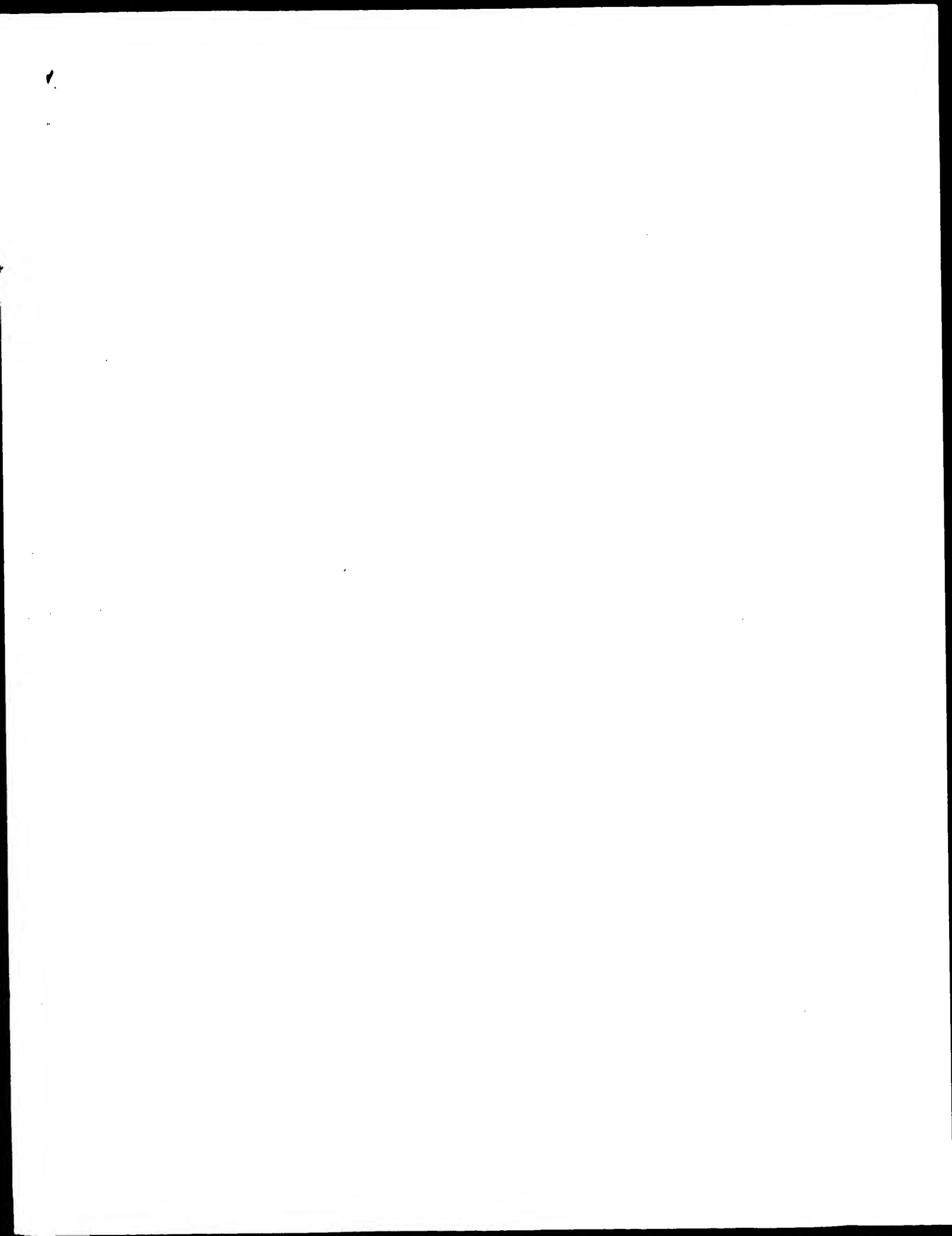
Database: p1r62  
 1:p1r1 2:p1r2 3:p1r3 4:p1r4

Statistics: Mean 28.844; Variance 40.415; scale 0.714  
 Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description	Pred. No.
No matches found.					

Search completed: Thu Jun 22 21:26:45 2000  
 Job time : 7 secs.



\*\*\*\*\*  
 MIPSEK  
 (TM)  
 \*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

Mparch<sub>LP</sub> protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Jun 22 21:25:48 2000; MasPar time 3.90 Seconds

Tabular output not generated. 195.186 Million cell updates/sec

Title: >US-09-332-866-1  
 Description: (1-25) from US09332866.pep  
 Perfect Score: 183  
 Sequence: 1 EEFLPKKLCQVDLHVISNDVCAQV 25

Scoring table: PAM 150  
 Gap 15

Searched: 83857 seqs, 30454973 residues

Post-processing: Minimum Match 0%  
 Listing first 45 summaries  
 Maximum DB seq length 25

Database: swiss-prot38  
 1:swissprot

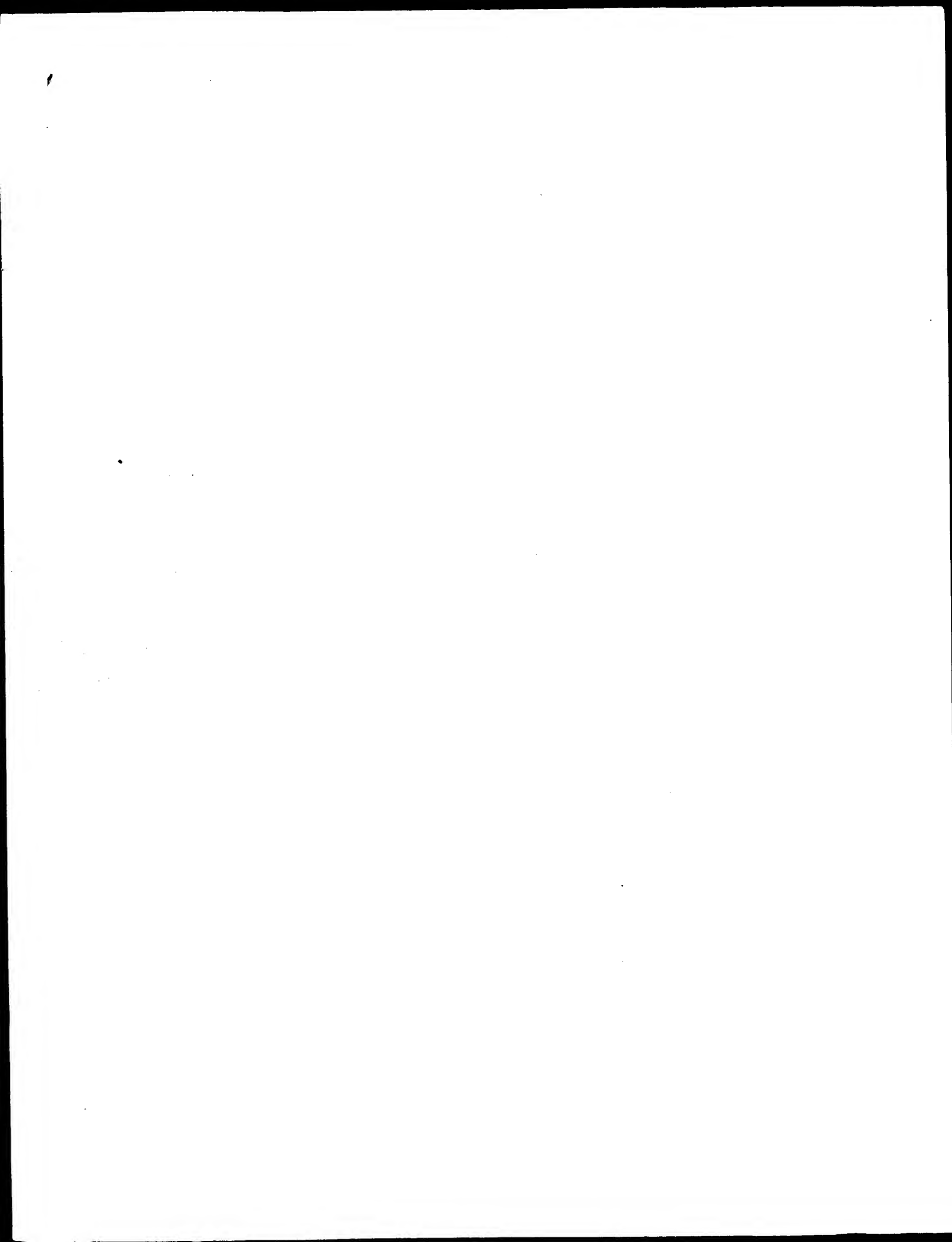
Statistics: Mean 29.602; Variance 37.132; scale 0.797

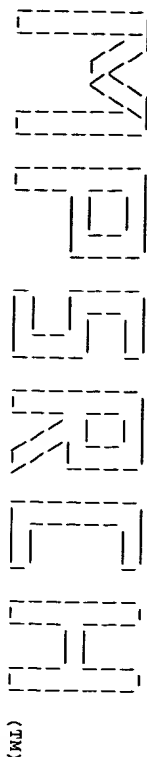
Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description	Pred. No.
No matches found.					

Search completed: Thu Jun 22 21:25:53 2000  
 Job time : 5 secs.





Release 3.1a John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

Mparch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Thu Jun 22 21:26:11 2000; Maspar time 9.41 Seconds  
184.120 Million cell updates/sec  
Tabular output not generated.

Title: >US-09-332-866-1  
Description: (1-25) from US09332866.pep  
Perfect Score: 183  
Sequence: 1 EEFLTPKKLQCVDLHVISNDVCAQV 25

Scoring table:  
PAM 150  
Gap 15

Searched: 225878 seqs, 69334122 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries  
Maximum DB seq length 25

Database: sptrembl12  
1:sp\_archaea 2:sp\_bacteria 3:sp\_fungi 4:sp\_human  
5:sp\_invertebrate 6:sp\_mammal 7:sp\_mhc 8:sp\_organelle  
9:sp\_phase 10:sp\_plant 11:sp\_rodent 12:sp\_unclassified  
13:sp\_vertebrate 14:sp\_virus

Statistics: Mean 28.558; Variance 35.870; scale 0.796

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description	Pred. No.
No matches found.					

Search completed: Thu Jun 22 21:26:22 2000  
Job time : 11 secs.





Transl. Seq 1 AA → 104 Size = 25

W T E S R E H (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_tpn n.a. - n.a. Smith-Waterman search, using a protein query  
which has been backtranslated into n.a. using IUPAC symbols

Run on: Fri Jun 23 17:30:31 2000; MasPar time 285.10 Seconds  
1121.254 Million cell updates/sec  
Tabular output not generated.

Title: >US-09-332-866-1  
Description: (1-25) from US09332866.pep  
Perfect Score: 316  
N.A. Sequence: 1 GARGARTTYTNACNCCNAA.....AYGAVGTNTGYGCNCARGTN 75  
Comp: CTCTYARARANTGNGGNTT.....TRCTRCANACRCGNGTTCAN

Scoring table: TABLE bktranslated2  
gap 40

Mismatch STD : Dbase 0; Query 0

Searched: 23132 seqs, 2311144889 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries  
Maximum DB seq length 25

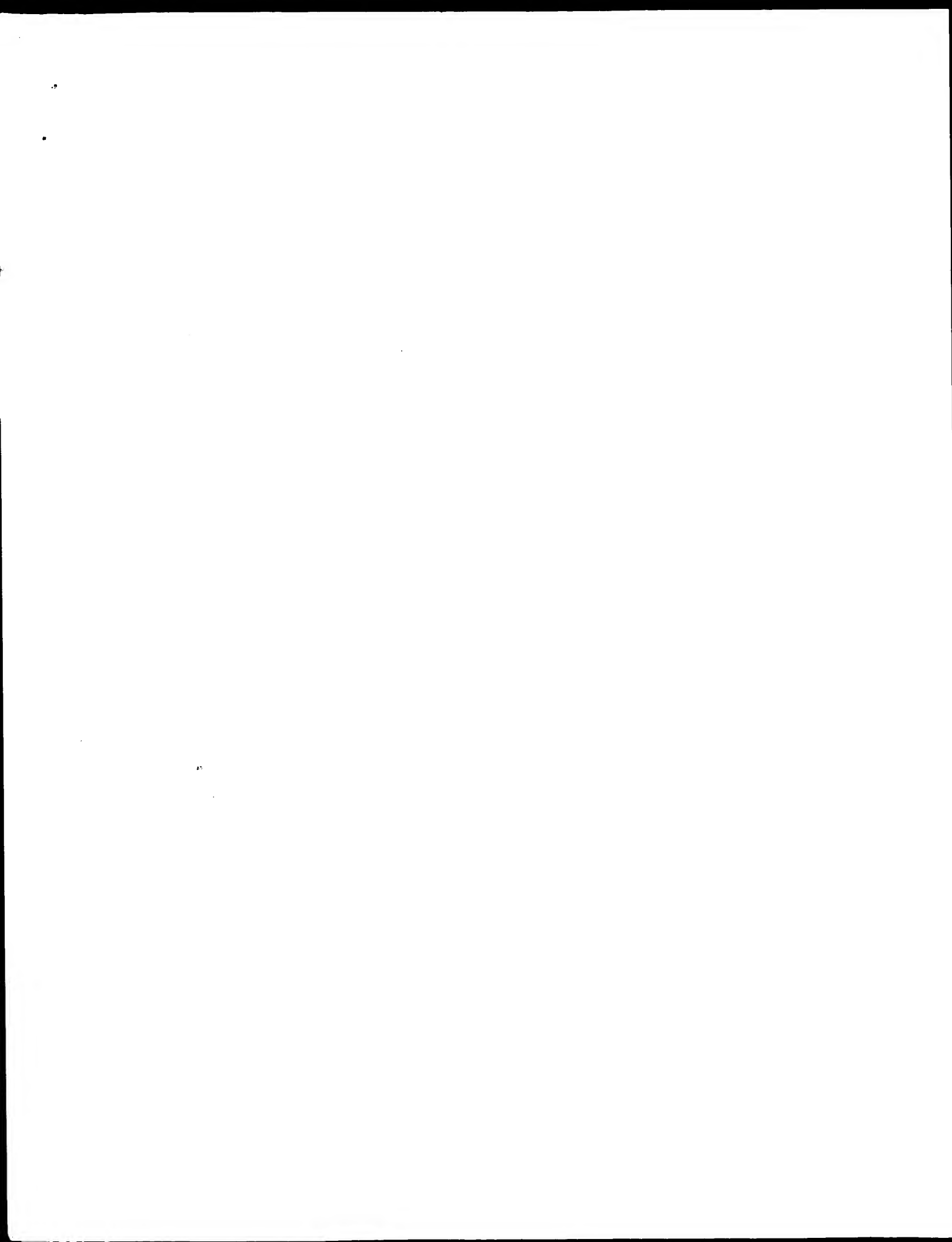
Database: emb161  
1:em\_ba1 2:em\_ba2 3:em\_fun 4:em\_htg1 5:em\_htg2 6:em\_htg3  
7:em\_hum1 8:em\_hum2 9:em\_hum3 10:em\_hum4 11:em\_in  
12:em\_om 13:em\_or 14:em\_ov 15:em\_pat 16:em\_pl 17:em\_ro  
18:em\_sts 19:em\_un 20:em\_v1  
Database: genbank116  
21:gb\_htg1 22:gb\_htg10 23:gb\_htg11 24:gb\_htg12  
25:gb\_htg13 26:gb\_htg14 27:gb\_htg2 28:gb\_htg3 29:gb\_htg4  
30:gb\_htg5 31:gb\_htg6 32:gb\_htg7 33:gb\_htg8 34:gb\_htg9

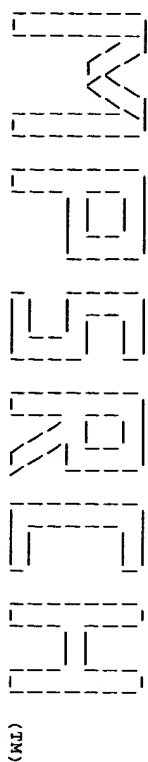
Statistics: Mean 78.372; Variance 327.890; scale 0.239  
Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
No matches found.						

Search completed: Fri Jun 23 17:35:29 2000  
Job time : 298 secs.





Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

Mpsrch\_tpn n.a. n.a. Smith-Waterman search, using a protein query  
which has been backtranslated into n.a. using IUPAC symbols

Run on: Fri Jun 23 17:24:02 2000; MasPar time 177.49 Seconds  
Tabular output not generated. 1034.812 Million cell updates/sec

Title: >US-09-332-866-1  
Description: (1-25) from US09332866.pep  
Perfect Score: 316  
N.A. Sequence: 1 GANGARTTYTNACNCCNAA.....AYGAGTNTGYGCNCARGTN 75  
Comp: CTYCTVARRANTGNGGNT.....TRCTRCANACRCNGGTTCAN

Scoring table: TABLE bkttranslate2  
Gap 40

Nmatch STD: Dbase 0; Query 0

Searched: 667327 seqs, 1224492533 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries  
Maximum DB seq length 25

Database: genbank2-116  
1:gb\_pat 2:gb\_ph 3:gb\_pl1 4:gb\_pl2 5:gb\_pl3 6:gb\_pr1  
7:gb\_pr2 8:gb\_pr3 9:gb\_pr4 10:gb\_pr5 11:gb\_ro 12:gb\_sts  
13:gb\_sy 14:gb\_un 15:gb\_v1

Statistics: Mean 49.932; Variance 184.198; scale 0.271

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Query Match	Length	DB ID	Description	Pred. No.
------------	-------------	--------	-------	-------------	-----------

No matches found.

Search completed: Fri Jun 23 17:27:04 2000  
Job time : 182 secs.



RELEASE

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

Mpsrch\_tpm n.a. - n.a. Smith-Waterman search, using a protein query  
which has been backtranslated into n.a. using IUPAC symbols

Run on: Fri Jun 23 18:53:33 2000; Maspar time 22.92 Seconds  
791.884 Million cell updates/sec  
Tabular output not generated.

Title: >US-09-332-866-1  
Description: (1-25) from US09332866.pep  
Perfect Score: 316  
N.A. Sequence: 1 GARGATTYYTNACNCNA.....AYGAYGTNGYGCNCARGTN 75  
Comp: CTYCTYAARRANTGNGNGNT.....TCTRCANACRCNGNTTCAN

Scoring table: TABLE bkttranslate2  
Gap 40

Mmatch STD : Dbase 0; Query 0

Searched: 311545 seqs, 121019393 bases x 2

Post-processing: Minimum Match 08  
Listing first 45 summaries  
Maximum DB seq length 25

Database: n-geneseq36  
1:geneseqn

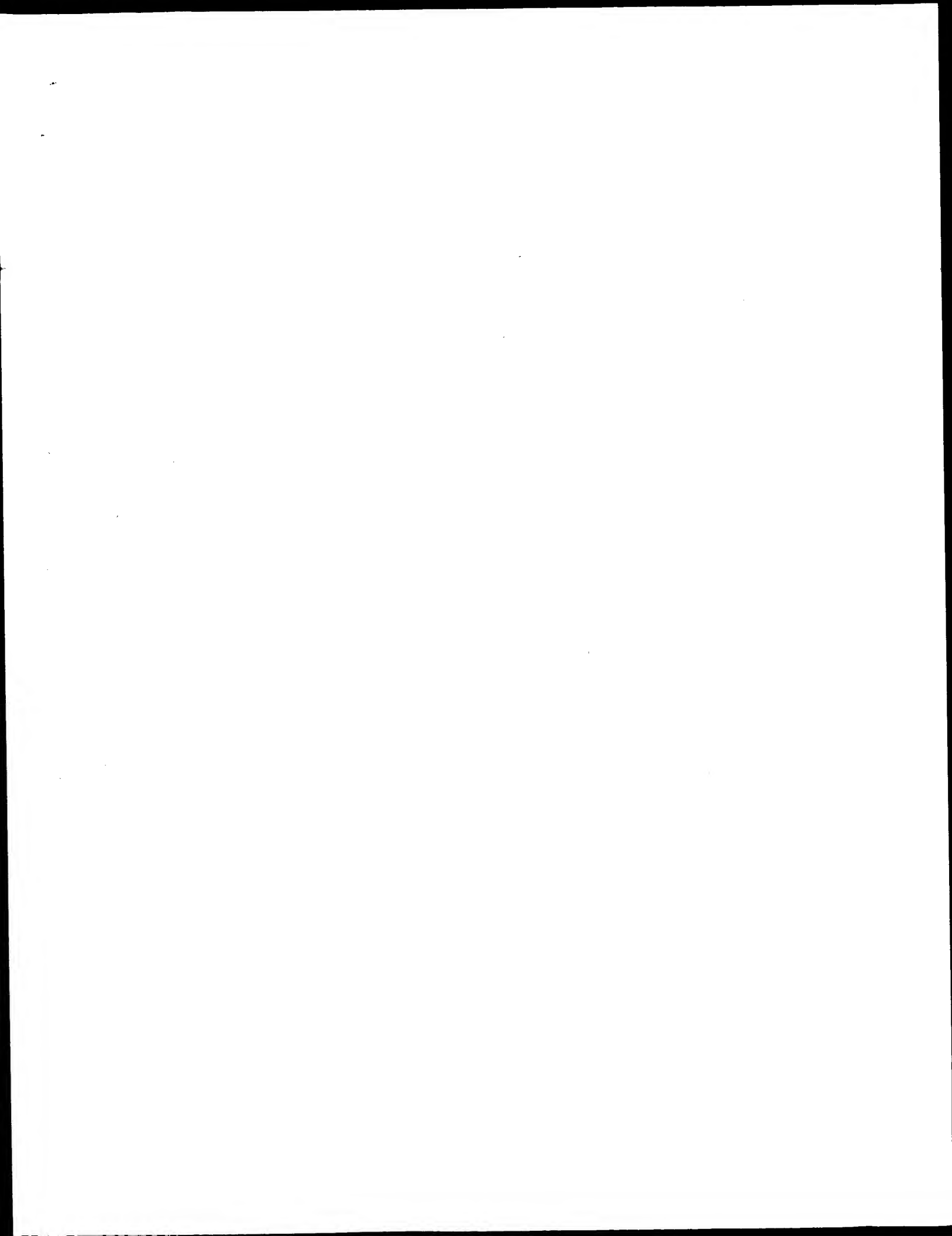
Statistics: Mean 40.189; Variance 176.566; scale 0.228

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	No.	Score	Match	Length	DB	ID	Description	Pred.	No.
No matches found.										

Search completed: Fri Jun 23 18:54:00 2000  
Job time : 27 secs.



WPISEIRH

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

Mpsrch\_tpn n.a. - n.a. Smith-Waterman search, using a protein query  
which has been backtranslated into n.a. using IUPAC symbols

Run on: Fri Jun 23 18:54:17 2000; Maspar time 15.21 seconds  
640.747 Million cell updates/sec  
Tabular output not generated.

Title: >US-09-332-866-1  
Description: (1-25) from US09332866.pep  
Perfect Score: 316  
N.A. Sequence: 1 GANGARTTYTNACNCNA.....AYGATGNTGYGNCARGTN 75  
Comp: CTYCTVAAARANTNGSGNT.....TRCTRCANACRCGNGTYCAN

Scoring table: TABLE dkttranslate2  
Gap 40

Nmatch STD : Dbase 0; Query 0

Searched: 230473 seqs, 64993002 bases x 2

Post-processing: Minimum Match 08  
Listing first 45 summaries  
Maximum DB seq length 25

Database: n-Issued  
1:5A\_COMB 2:5B\_COMB 3:5C\_COMB 4:5D\_COMB 5:6\_COMB  
6:PCT\_COMB 7:backfiles1

Statistics: Mean 37.556; Variance 145.681; scale 0.258

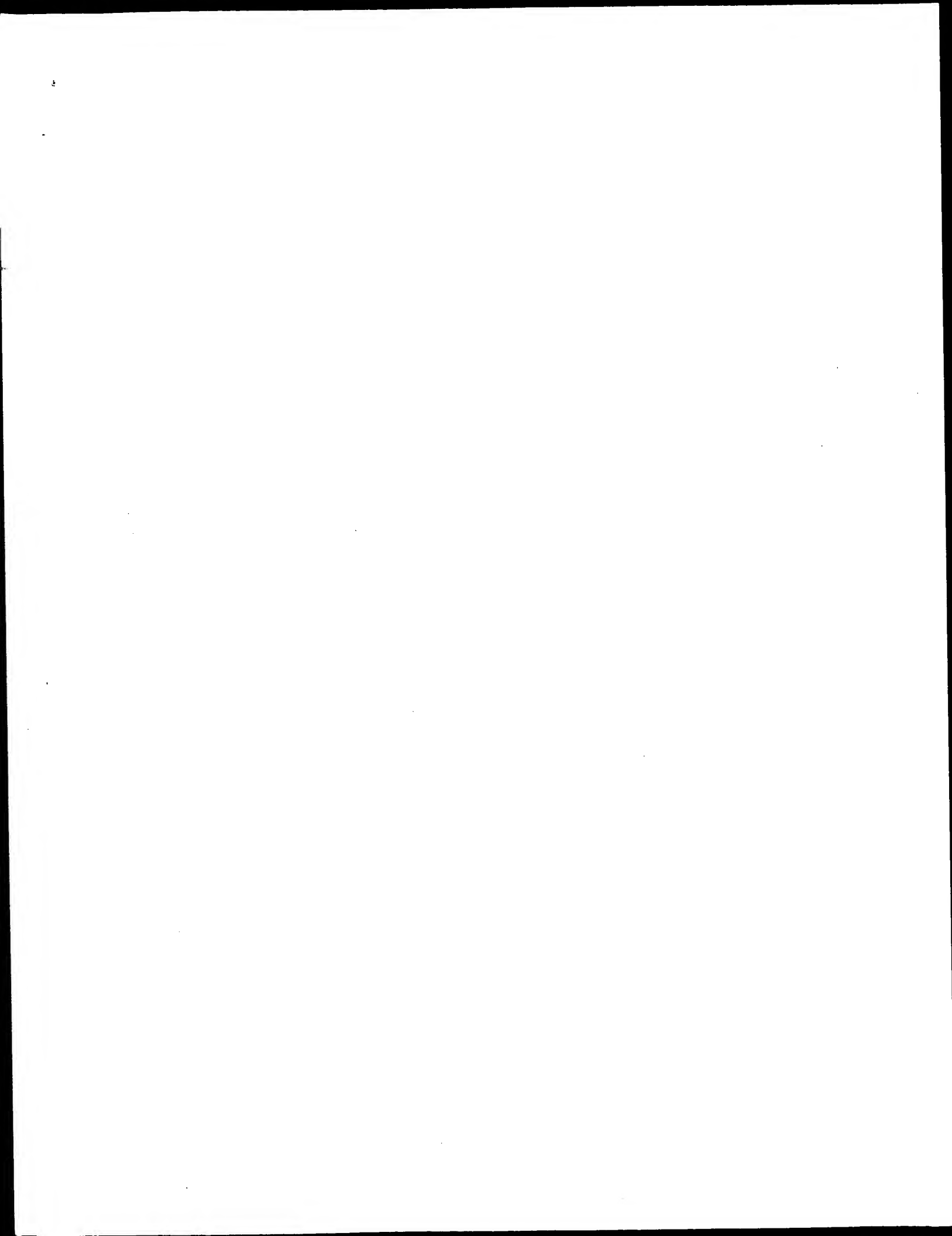
Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	No.	Score	Match	length	DB	ID	Description	Pred. No.
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

No matches found.

Search completed: Fri Jun 23 18:54:36 2000  
Job time : 19 secs.







Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_tpn n.a. - n.a. Smith-Waterman search, using a protein query  
which has been backtranslated into n.a. using IUPAC symbols

Run on: Fri Jun 23 18:54:56 2000; MasPar time 248.75 Seconds  
809.968 Million cell updates/sec  
Tabular output not generated.

Title: >US-09-332-866-1  
Description: (1-25) from US09332866.pep  
Perfect Score: 316  
N.A. Sequence: 1 GARGARTTYTNACNCNNA.....AYGAVGNTGCGNCARGTN 75  
Comp: CTCTTAAARANTGNGNTT.....TRCTRCANACRCGNGTYCAN

Scoring table: TABLE bktranslate2  
Gap 40

Nmatch STD : Dbase 0; Query 0

Searched: 3601417 seqs, 1343217448 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries  
Maximum DB seq length 25

Database: n-pending1  
1:U6 2:U7 3:U80 4:U81A 5:U81B 6:U81C 7:U81D 8:U82A  
9:U82B 10:U82C 11:U83A 12:U83B 13:U84A 14:U84B 15:U84C  
16:U85 17:U86 18:U87A 19:U87B 20:U87C 21:U88A 22:U88B  
23:U88C 24:U88D 25:U89A 26:U89B 27:U89C 28:U89D 29:U89E  
30:U89F 31:U89G 32:U89H 33:U89I 34:U90A 35:U90B 36:U90C  
37:U90D 38:U90E 39:U91A 40:U91B 41:U91C 42:U92A 43:U92B  
44:U92C 45:U92D 46:U92E 47:U93A 48:U93B 49:U93C 50:U93D  
51:U93E 52:U93F 53:U94A 54:U94B 55:U94C 56:U95 57:NEW08  
58:NEW09

Statistics: Mean 49.471; Variance 110.484; scale 0.448

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query Match	Score	Length	ID	Description	Pred. No.
No matches found.						

Search completed: Fri Jun 23 18:59:11 2000  
Job time : 255 secs.



\*\*\*\*\*  
 WIDEORIT  
 (TM)  
 \*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

Mparch\_tpn n.a. - n.a. Smith-Waterman search, using a protein query  
 which has been backtranslated into n.a. using IUPAC symbols

Run on: Fri Jun 23 17:27:21 2000; MasPar time 166.66 Seconds  
 891.071 Million cell updates/sec  
 Tabular output not generated.

Title: >US-09-332-866-1  
 Description: (1-25) from US09332866.pep  
 Perfect Score: 316  
 N.A. Sequence: 1 GARGARTYYTNACGCCNA.....AYGAVGTNTGYGCNCARGTN 75  
 Comp: CTCTCYARRRANTGNGNTT.....TRCTRCANACRCGNGTYCAN

Scoring table: TABLE bkttranslated2  
 Gap 40

Nmatch STD : Dbase 0: Query 0

Searched: 1741034 segs, 990042266 bases x 2

Post-processing: Minimum Match 0%  
 Listing first 45 summaries  
 Maximum DB seq length 25

Database: n-pending2  
 1:PCT 2:U6000 3:U6001A 4:U6001B 5:U6002A 6:U6002B  
 7:U6002C 8:U6003A 9:U6003B 10:U6004A 11:U6004B 12:U6004C  
 13:U6005 14:U6006 15:U6007 16:U6008A 17:U6008B 18:U6008C  
 19:U6009A 20:U6009B 21:U6010A 22:U6010B 23:U6011A  
 24:U6011B 25:U6012A 26:U6012B 27:U6013A 28:U6013B  
 29:U6014A 30:U6014B 31:U6014C 32:U6015 33:U6016A  
 34:U6016B 35:U6016C 36:U6017A 37:U6017B 38:U6017C  
 39:U6018A 40:U6018B 41:U6018C 42:U6019 43:NEWP 44:NEWU60  
 45:NEWU60

Statistics: Mean 51.092; Variance 101.380; scale 0.504

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

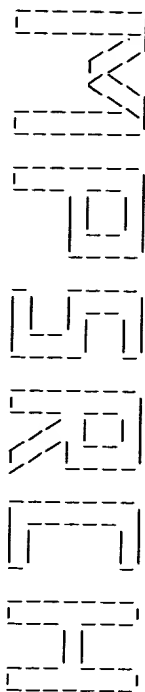
# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
*****							

No matches found.

Search completed: Fri Jun 23 17:30:12 2000  
 Job time : 171 secs.





(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

Mpsrch\_tpn n.a. - n.a. Smith-Waterman search, using a protein query  
which has been backtranslated into n.a. using IUPAC symbols

Run on: Fri Jun 23 17:35:51 2000; Maspar time 4610.56 seconds  
Tabular output not generated. 65.934 Million cell updates/sec

Title: >US-09-332-866-1  
Description: (1-25) from US09332866.pep  
Perfect Score: 316  
N.A. Sequence: 1 GARGARTTYTNACNCCNNA.....AYGAYGTNTGCGNCARGTN 75  
Comp: CTYCTYAARRANTNGNGNTT.....TCTRCANACRCNGNGTYCAN

Scoring table: TABLE bkttranslate2  
Gap 40

Mmatch STD : Dbase 0; Query 0

Searched: 4857316 seqs, 2026511650 bases x 2

Post-processing: Minimum Match 08

Listing first 45 summaries  
Maximum DB seq length 25

Database: emb1-est61  
1:em\_est11 2:em\_est12 3:em\_est13 4:em\_est14 5:em\_est15  
6:em\_est20 7:em\_est21 8:em\_est22 9:em\_est23 10:em\_est29  
11:em\_est3 12:em\_est30 13:em\_est31 14:em\_est5 15:em\_gss1  
16:em\_gss12 17:em\_gss2 18:em\_gss3 19:em\_gss4 20:em\_gss5  
21:em\_gss6  
genbank-est116  
22:gb\_est1 23:gb\_est10 24:gb\_est11 25:gb\_est12  
26:gb\_est13 27:gb\_est14 28:gb\_est15 29:gb\_est16  
30:gb\_est17 31:gb\_est18 32:gb\_est19 33:gb\_est2  
34:gb\_est20 35:gb\_est21 36:gb\_est22 37:gb\_est23  
38:gb\_est24 39:gb\_est25 40:gb\_est26 41:gb\_est27  
42:gb\_est28 43:gb\_est29 44:gb\_est3 45:gb\_est30  
46:gb\_est31 47:gb\_est32 48:gb\_est33 49:gb\_est34  
50:gb\_est35 51:gb\_est36 52:gb\_est37 53:gb\_est38  
54:gb\_est39 55:gb\_est4 56:gb\_est40 57:gb\_est41  
58:gb\_est42 59:gb\_est43 60:gb\_est44 61:gb\_est45  
62:gb\_est46 63:gb\_est47 64:gb\_est5 65:gb\_est6 66:gb\_est7  
67:gb\_est8 68:gb\_est9 69:gb\_gss1 70:gb\_gss10 71:gb\_gss11  
72:gb\_gss12 73:gb\_gss13 74:gb\_gss14 75:gb\_gss15  
76:gb\_gss16 77:gb\_gss2 78:gb\_gss3 79:gb\_gss4 80:gb\_gss5  
81:gb\_gss6 82:gb\_gss7 83:gb\_gss8 84:gb\_gss9

Statistics: Mean 53.419; Variance 70.507; scale 0.758

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

No. Score Match Length DB ID Description Pred. No.  
-----

No matches found.  
Search completed: Fri Jun 23 18:53:14 2000  
Job time : 4643 secs.

Result Query

SUMMARIES

